A landscape genetics perspective on the spatial dynamics of hybridization between two species of wall lizards

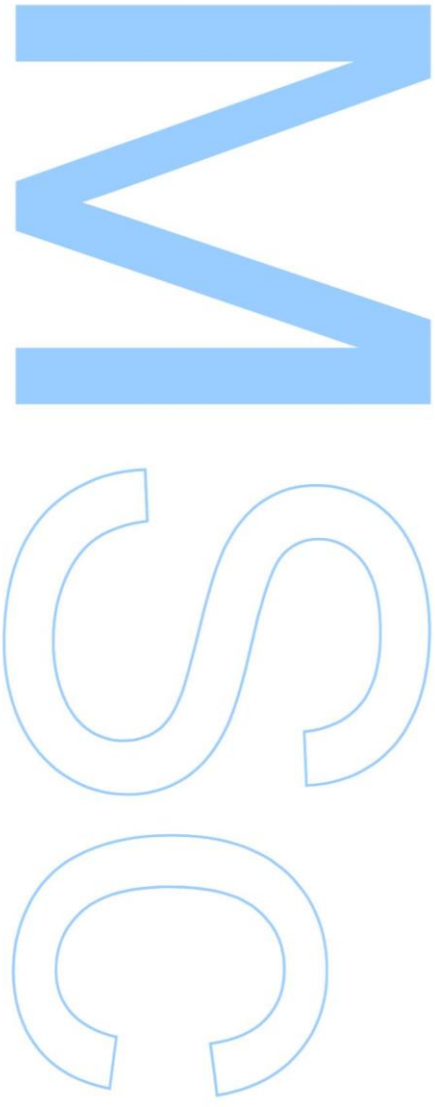
Mariana Jorge Meneses Correia Ribeiro
Mestrado em Biodiversidade, Genética e Evolução
Departamento de Biologia
2014

Orientador
Catarina Pinho, Post-Doc, Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), Universidade do Porto

Co-orientador
Antigoni Kaliontzopoulou, Post-Doc, Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), Universidade do Porto

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, _____/_____/_____



“We used to make fun of Edgar Anderson by saying that he was finding hybrids under every bush. Then we realized that even the bushes were hybrids”

Warren H. Wagner

Agradecimentos

E agora que mais uma fase de aprendizagem termina para dar lugar a muitas outras, não posso deixar de agradecer a todos aqueles que me ajudaram nesta jornada.

Em primeiro lugar quero agradecer à minha orientadora, Catarina Pinho por todo o apoio, por tudo o que me ensinou e principalmente pela paciência perante todas as perguntas e dúvidas de quem está a iniciar um percurso em ciência. Especialmente, quero agradecer por sempre ter ouvido as minhas opiniões e ideias com todo o respeito e sempre com o sorriso e simpatia que te caracterizam. Obrigada.

À Antigoni por estar sempre presente e me ouvir mesmo quando o assunto não era a sua especialidade. Principalmente por estar sempre a apoiar-me quando via que estava mais em baixo. Quero que saibas que as tuas palavras de apoio, força e reconhecimento foram sempre essenciais para conseguir levar este trabalho até ao fim. Nunca vou esquecer cada ensinamento, cada palavra amiga. σας ευχαριστώ.

Ao Miguel Ángel Carretero pelos conhecimentos incríveis sobre lagartixas, e principalmente por encontrar sempre mil e uma explicações biológicas para explicar os resultados. Além disso, agradeço todo o entusiasmo que demonstrou ao longo do trabalho o que me fez gostar cada vez mais de lagartixas. Sei que ainda tenho muito mais para apreender contigo. Gracias.

À Xana Sá Pinto que apesar de forma não-oficial ajudou na minha orientação durante todo o trabalho. Obrigada pelas perguntas difíceis que me fizeram ter vontade de arrancar cabelos mas que ao mesmo tempo me ensinaram não a decorar conceitos mas a percebê-los e aplicá-los. Sei que tudo o que apreendi contigo vai ser muito importante ao longo da minha futura carreira. Obrigada por todo o teu apoio.

À Susana Lopes…não há palavras. Se hoje tenho resultados genéticos para apresentar a ti se deve. Obrigada por nunca desistires durante meses de optimização. Mas mais importante que isso, obrigada pela amizade e por me aturares sempre que estava chateada. Nunca me vou esquecer de ti.

Ao Guilherme Dias, por estar sempre comigo a apanhar aquelas lagartixas demoníacas, por me ensinar como se fazia uma extração e um PCR, pelos almoços

descontraídos e cafezinhos nas horas de mais calor. Acima de tudo, obrigada pelas conversas e amizade.

A todo o pessoal do CTM pela ajuda no laboratório com especial destaque à Sofia Mourão por responder sempre com um sorriso a todas as perguntas básicas e à Patrícia Ferreira por conseguir sempre encaixar as minhas amostras e simplesmente pela pessoa que é.

Aos meus colegas de mestrado pelos momentos de loucura, pelas gargalhadas ao almoço e pelo companheirismo sempre presente. Este ano sem vocês não tinha sido o mesmo. Em especial à Isabel por ser a louca que é e por me fazer sempre rir com as barbaridades que diz. Obrigada amiga, foste essencial neste ano e no resto da minha vida; À Ana Cristina por estar sempre lá quando foi preciso e pelas experiências partilhadas; À Sara por me fazer sempre rir cada vez que ia a internet; Ao Zé pela calma irritante de quem não tem uma tese para entregar; Ao Pedro por me ter proporcionado momentos hilariantes enquanto criava aquele gif épico.

À Mafalda por conseguir ser tão louca como eu e por ser a minha companheira na palhaçada e brincadeira. Juntas somos maléficas!; À Luana pelas novas *skills* no paint, por toda a companhia, conversas, ajuda sempre que foi necessário (sem ti esta tese não era entregue). Sei que vou levar uma amiga para a vida; À Carolina…não há como descrever este ano querida…conhecer-te melhor foi das melhores coisas que me aconteceu. Meninas, sempre juntas na alegria e na desgraça!

Aos meus amigos pelos momentos de descontração e por toda a ajuda sempre que foi necessário.

Ao Miguel por ter aguentado as minhas ausências sem reclamar. Por me ouvir falar de coisas que não entende. Enfim…obrigada por estares na minha vida.

Aos meus pais e ao meu padrasto por me terem aturado quando os dias no laboratório corriam mal. Em especial a minha mãe por todas as lancheiras a rebentar de comida e por ter me dizer sempre para te calma que tudo se ia resolver. Aos meus irmãos e sobrinhos por compreenderem esta fase mais difícil com o mesmo amor de sempre.

Por último, ao Blacky por estar sempre sentado ao meu lado nas longas horas de escrita.

Sumário

Compreender o isolamento reprodutivo e as barreiras ao fluxo génico é essencial para entender a especiação. Desta forma, as zonas híbridas, áreas onde espécies geneticamente diferentes entram em contacto e se cruzam, permitem o estudo dos processos envolvidos no isolamento reprodutor, permitindo identificar potenciais barreiras ao fluxo génico, estudar os padrões genómicos da seleção e introgressão e identificar os caracteres fenotípicos sob seleção. *Podarcis bocagei* e *P. carbonelli* constituem um bom modelo de estudo para a especiação uma vez que estas hibridam, mantendo no entanto a sua identidade genética e morfológica na sua restrita área de contacto. Neste projeto propusemo-nos a estudar a zona híbrida entre as duas espécies utilizando uma estratégia multidisciplinar. Especificamente, combinámos análises genéticas e morfológicas com o mapeamento geográfico a fina escala de cada indivíduo, de forma a compreender a dinâmica espacial desta zona.

Para a análise molecular utilizou-se o DNA mitocondrial e um conjunto de microssatélites para estimar o nível de miscigenação de cada indivíduo. Estes marcadores moleculares, confirmam o carácter bimodal desta zona híbrida reportado em estudos anteriores, permitindo a identificação de 144 indivíduos de *P. carbonelli*, 33 de *P. bocagei* e 18 híbridos, resultando numa proporção de híbridos de 9.2%. Foi possível identificar diferentes tipos de híbridos incluindo F1, F2 e retrocruzamentos com as classes parentais. Apesar da maior abundância de *P. carbonelli* na amostragem, os híbridos apresentam, maioritariamente, DNA mitocondrial de *P. bocagei*, sugerindo a ocorrência preferencial de cruzamentos interespecíficos entre fêmeas de *P. bocagei* e machos de *P. carbonelli*. Os nossos resultados parecem suportar a regra de Haldane sugerindo uma redução da fertilidade das fêmeas embora não suporte a sua inviabilidade. Estes resultados sugerem a existência de fortes barreiras ao fluxo génico na zona híbrida.

Além disso, foi utilizada uma combinação de caracteres folidóticos para analisar os possíveis efeitos da hibridação na morfologia das espécies e nas características morfológicas dos híbridos. *P. bocagei* e *P. carbonelli* são mais similares morfologicamente em zonas de simpatria do que em zonas alopátricas, não suportando a ocorrência de *reinforcement* a atuar sobre estas características morfológicas. A maior semelhança morfológica entre as duas espécies na zona de contacto pode estar relacionada com a ocorrência de introgressão devido ao processo de hibridação. Os indivíduos híbridos não apresentam diferenças significativas na morfologia em relação às formas parentais e não tendo sido detetadas formas

morfologicamente intermédias. Adicionalmente, analisamos os níveis de assimetria de diversos caracteres bilaterais para testar se a hibridação aumenta a instabilidade no desenvolvimento, com potenciais efeitos no *fitness* dos híbridos. Estas análises não revelaram diferenças significativas na instabilidade do desenvolvimento entre híbridos e formas parentais. Em consonância com estudos anteriores o presente estudo sugere que a hibridação entre as duas espécies de lagartixas analisadas é um fenómeno raro, embora constante nesta zona híbrida.

Dado que fortes barreiras ao fluxo génico foram detetadas, estudámos a dinâmica espacial desta zona híbrida à micro-escala geográfica, de forma a obter mais informação sobre os fatores que contribuem para limitar a hibridação entre estas espécies. Análises de clinos e interpolações mostram uma transição abrupta nos caracteres genéticos. Por outo lado, os caracteres morfológicos apresentam uma distribuição em forma de mosaico na zona híbrida que está provavelmente relacionada com a intra e inter-variabilidade existente nas espécies modelo. Os resultados sugerem a existência de uma barreira geográfica que limita o contacto entre as duas espécies. No entanto, a baixa frequência híbridos, quando comparados com indivíduos puros, sugere fortemente a existência de barreiras reprodutivas entre estas espécies. Mais estudos são necessários para identificar estas barreiras, e compreender a sua natureza e o seu papel na restrição do fluxo génico entre as espécies.

Neste estudo utilizou-se uma abordagem multidisciplinar, combinando dados genéticos, morfológicos e espaciais para analisar esta zona híbrida, de forma a relacionar os resultados com explicações biológicas. Esta abordagem permitiu caracterizar a zona híbrida e forneceu informação importante para compreender de que forma as barreiras geográficas e reprodutivas contribuem para limitar o fluxo génico entre espécies de *Podarcis*. Assim, salientamos a importância de uma abordagem multidisciplinar para investigar os processos que responsáveis pela formação de novas espécies e manutenção da sua integridade.

Palavras-chave: *P. bocagei*, *P. carbonelli*, zonas híbridas, barreiras reprodutivas, padrões espaciais.

Abstract

Understanding reproductive isolation and barriers to gene flow is essential to understand speciation. As such, hybrid zones, areas where genetically differentiated species come into contact and interbreed to some extent, allow the study of processes involved in the emergence of reproductive barriers, allowing to identify potential barriers to gene flow, to study the genomic patterns of selection and introgression, and to identify the phenotypic characters under selection. *Podarcis bocagei* and *P. carbonelli* are two lacertid lizards, endemic to the Iberian Peninsula. They provide a good model system to study speciation because they are known to interbreed in the wild to some extent, but also to maintain their genetic and morphological identities in their small contact area. In this work we studied the hybrid zone between these two species using a multidisciplinary approach. Specifically, we combined population genetics and morphological analyses with a fine-scale geographic mapping of individuals to understand the spatial dynamics of this hybrid zone.

For molecular analyses we used mitochondrial DNA and a combination of microsatellites to evaluate the level of admixture of each individual. These molecular markers revealed a strongly bimodal hybrid zone and allowed the identification of 144 individuals of *P. carbonelli* and 33 of *P. bocagei* as well as 18 hybrids, resulting in a proportion of hybrids of 9.2%. Among these, we found hybrids of multiple classes (F1, F2, backcrosses with each parental form and other types of hybrids). Moreover, despite the presence of more *P. carbonelli* individuals in our sample, we found that the majority of hybrids had a *P. bocagei* maternal origin, indicating a directionality of crosses where females of *P. bocagei* breed with males of *P. carbonelli* and hybrid males. Our results provide some support to Haldane's rule suggesting the reduced hybrid female fertility, although not complete unviability. These results suggest strong barriers to gene flow in the hybrid zone.

Further, we used a combination of pholidotic (scalation) traits to examine the effects of hybridization on the morphology of the species by exploring the morphological properties of hybrids. Morphologically, *P. bocagei* and *P. carbonelli* are more similar in sympatric than in allopatric areas, allowing us to reject reinforcement acting at these morphological traits. The higher similarity between the two species at the contact zone can be related with introgression due to long-term hybridization. Hybrid individuals do not differ in morphology from the parental forms while no morphologically intermediate forms were identified. No relation was detected between genetic and morphological assignments of hybrid individuals. Therefore, morphological

traits do not follow genetic admixture patterns. Also, we examined the level of asymmetry of several bilateral traits to test whether hybridization could be associated to increased developmental instability with potential effects on the fitness of hybrids. In this case, no differences in developmental stability were detected between hybrids and parental forms. Together with a previous study, our results suggest that although rare, hybridization has been regularly occurring at this hybrid zone.

As strong barriers to gene flow were suggested, we studied the spatial dynamics of the hybrid zone at the microgeographic scale, in order to obtain more insights on the factors limiting gene flow between these two wall lizards. A steep transition across the landscape was detected in genetic traits trough clinal analysis. On the other hand, morphological traits show a mosaic distribution across the area, possibly due to the high intra and interspecific morphological variability of the model species. Our results suggest a strong influence of a geographic barrier that seems to limit the contact between the two species. However, the low number of individuals with intermediate genotypes when compared to those with carrying pure genotypes, strongly suggest the existence of reproductive barriers between the two species. Further studies are required to identify these barriers and study its nature and its role impeding gene flow between these species.

In the present study, the combination of genetic, morphological and spatial evidence allowed to characterise the features of this interesting hybrid zone providing important information that allowed us to better understand how geographic and reproductive barriers interplay and limit gene flow between these two *Podarcis* species. This study also highlights the importance of multidisciplinary approaches in hybrid zone research and how these approaches allow to investigate how species integrity is maintained, to enhance biological interpretation of the results, eventually leading to a better understanding of the processes driving speciation.

Key words: *P. bocagei*, *P. carbonelli*, hybrid zones, reproductive barriers, spatial patterns.

Index

List of Figures

List of Tables

List of Abbreviations

HW – Hardy-Weinberg
LD – Linkage disequilibrium
LE – Linkage equilibrium
SVL – Snout-vent length
DNA – Deoxyribonucleic acid
mtDNA – Mitochondrial DNA
PCR – Polymerase chain reaction
rRna – ribosomal RNA
bp – Base pair
Ho – Observed heterozigosity
He – Expected heterozigosity
Na – Number of alleles
Pa – Private alleles
AMOVA – Analyses of Molecular Variance
Fst – Fixation index
FCA – Factorial Correspondence Analysis
K – Number of clusters
MCMC – Markov Chain Monte Carlo
qi – Admixture proportion
PP – Posterior probability
Pid – Probability of identity
PidSib – Probability of identity in siblings
CSN – Colaria scales
FPN – Femoral pores
GSN – Gularia scales
SCGN – Supraciliary granules
SDLN – Subdigital lamellae
STSN – Supratemporal scales
VSN – Ventral scales
SCSN – Supraciliary scales
SLSN – Supralabial scales
IN_F – Contact between the internasal and frontal scales
O_IP – Contact between the occipital and interparietal scales
MASS – Presence of the masseteric scale

TYMP – Presence of the tympanic scale

ANOVA – Analysis of Variance

PCA – Principal Component Analysis

DA – Discriminant Analysis

FA – Fluctuating asymmetry

AI – Asymmetry index

PCoA – Principal Coordinate Analysis

IDW – Inverse distance-weighted

ED – Euclidian distances

AICc – Akaike's information criterion

SD – Standard deviation

1. Introduction

1.1. Speciation

Biologists are trying to understand the processes involved in species differentiation since before Darwin (Coyne & Orr 2004; Sobel *et al.* 2010). Under the biological species concept (Mayr 1942, 1963), speciation can be defined as the development of mechanisms responsible for reproductive isolation (Turelli *et al.* 2001; Wiens 2004; Butlin *et al.* 2012). Therefore, understanding speciation becomes the study of when, where and how gene flow breaks down. Speciation is a complex process because it integrates several mechanisms such as ecology, behaviour, morphology and interaction between multilocus genotypes. The study of speciation is based on the clarification of the mechanisms that create reproductive isolation. Therefore, understanding the genetics and genomics of speciation and connecting speciation processes is important to the research on biodiversity patterns (Butlin *et al.* 2012).

Barriers to gene exchange may evolve during long periods of spatial isolation (Abbott *et al.* 2013). However, it is also possible that incipient species contact before developing complete reproductive barriers. An extreme example is divergence without any spatial isolation, mediated exclusively by divergent selection, but intermediate scenarios, such as contact during variable periods because of e.g. range changes may be more common (Abbott *et al.* 2013). The study of reproductive barriers in areas where incompletely isolated species contact is thus essential to understand speciation and to evaluate the relative importance of such isolating mechanisms. Typically, barriers to gene exchange are classified with respect to whether they act before ("prezygotic") or after ("postzygotic") hybrid zygote formation (Butlin *et al.* 2012). Prezygotic isolation mechanisms, normally based on assortative mating (but also in other processes such as sperm competition), have been considered more important barriers to gene flow, potentially having a higher influence in the process of speciation, than postzygotic barriers, which are not always detected (Sobel *et al.* 2010).

In species differentiation, normally the unit of adaptation is the gene or a set of interacting genes, which means that the genome does not work as a whole and the levels of gene flow and differentiation vary along the genome (Wu 2001). The genetic basis of reproductive isolation is usually thought to result from epistatic incompatibilities

in genes responsible for physiological, anatomical, ecological or behavioural isolation (Turelli *et al.* 2001; Coyne & Orr 2004). Gene flow becomes more difficult as the number of genes that present incompatibilities between diverging taxa increase, due to the co-adaptation of different loci within populations (Orr 1995). To understand speciation it is necessary to know how genetic differences between species evolve and to determine the effect of these differences on gene flow (Barton & Bengtsson 1986). When divergent selection is stronger than gene flow, reproductive barriers between different taxa appear (Kisel & Barraclough 2010). Also, strong selection on only one or few traits could be more effective in promoting adaptive divergence with gene flow, and weaker selection on multiple traits creates a more general barrier to gene flow (Nosil 2008). Therefore, significant reproductive isolation only appears when genetic differences cause a substantial reduction in gene flow at the majority of loci, or if they promote the evolution of further differences that can lead to a strong barrier to gene exchange (Barton & Bengtsson 1986). The study of speciation is based on the clarification of the mechanisms that create reproductive isolation. These mechanisms can be addressed when reproductive barriers between species are not complete, thus when gene flow exists and species hybridize.

1.2. Hybridization: definition and consequences

Hybridization is defined as "the interbreeding of individuals from two populations, or groups of populations, which are distinguishable on the basis of one or more heritable characters" (Harrison 1993). The different groups that hybridize could be species, subspecies or races that are different from each other on the basis of various characters such as morphology (e.g. body colour, meristic traits such as the number of scales in lizards), behaviour (e.g. mate choice), cytological (e.g. chromosome numbers) and molecular markers (e.g. DNA sequences) (Kawakami & Butlin 2001).

Natural hybridization and gene flow were, and still are, controversial amongst taxonomists because they violate the strict biological species concept (Mayr 1942, 1963). Traditionally, these processes were considered rare in the wild and they were relatively neglected, not being considered an important evolutionary process, particularly in animals (Arnold 1992; Mallet 2005). This view has changed thanks to new cytogenetic studies and molecular data. Nowadays, natural hybridization is known to occur much more frequently than initially anticipated. In fact, hybridization between

species has been estimated to occur in 10% of animal species and 25% of plants (Mallet 2005). Moreover, hybridization is now recognized as an important evolutionary phenomenon, with variable effects on the differentiation of populations. On one hand, hybridization may slow or reverse differentiation between incipient species by allowing gene flow and recombination to break apart co-adapted gene arrangements (Mallet 2007; Abbott *et al.* 2013). In some cases hybridization may lead to the reduction or even extinction of rare species by genetic homogenization or disruption of local adaptation (Allendorf *et al.* 2001). On the other hand, it has long been recognized that hybridization has the potential to create adaptive variation, functional novelty and even lead to the formation of new species (Lewontin & Birch 1966; Seehausen 2004). Indeed, it has been shown that hybridization can accelerate speciation via adaptive introgression or cause near-instantaneous speciation by allopolyploidization (Abbott *et al.* 2013). It can also lead to speciation if the new hybrid lineages become reproductively isolated from the parental forms, a process known as hybrid speciation (Mallet 2007; Abbott *et al.* 2013). This can happen e.g. when a hybrid recombinant genome is as fit as or fitter than the parental forms in areas of intermediate habitat (Dobzhansky 1940; Kawakami & Butlin 2001).

The frequency of hybridization is affected by a series of factors, such as the reproductive barriers (either prezygotic or postzygotic) that decrease the frequency of hybrids in the wild. Time since divergence is also very relevant, since more closely related species have the tendency to hybridize more often than species that diverged longer ago and which have had more time to develop mechanisms of reproductive isolation and accumulate incompatibilities. Finally, there are taxon-specific differences and some groups hybridize more often than others (Mallet 2005). The relative contribution of different factors in determining the frequency of hybridization can be efficiently investigated in hybrid zones.

1.2.3. Hybrid zones

Hybrid zones have received extensive attention and have been referred to as important “windows on the evolutionary process” (Harrison 1990). Hybrid zones are geographically narrow contact zones where distinct groups of individuals meet, mate and produce hybrids (Barton & Hewitt 1985). They have been observed in a wide variety of organisms (Barton & Hewitt 1985; Kawakami & Butlin 2001; Abbott *et al.* 2013). Normally, hybrid zones present a mixture of genotypes ranging from first generation hybrids to individuals with repeatedly admixed genomes that result from

multiple generations of backcrossing (Kawakami & Butlin 2001). This leads to introgression, the acquisition by one species of genetic material from another (Baack & Rieseberg 2007). The extent of introgression is highly dependent on the effects of fitness (if hybrids present lower fitness when compared to parental forms introgression will be minimum and vice-versa) and reproductive isolation, as well as on genetic linkage relationships (Barton 1979; Wu 2001; Gompert *et al.* 2012). Due to the typically strong assortative mating between closely related species that live in sympatry, first generation hybrids are often difficult to produce in animals with internal fecundation (Mallet 2005). However, after the appearance of this class of hybrids, the backcrossing to the parental forms is much more straightforward if the hybrid individuals are viable and fertile (Mallet 2005, 2008).

Hybrid zones can be classified according to the distribution of genotypic classes. The unimodal hybrid zone is based on a hybrid swarm where intermediate hybrid genotypes predominate (Jiggins & Mallet 2000). A trimodal hybrid zone is characterized by a defined group of hybrids with intermediate allele frequencies and extensive genotyping variance caused by several levels of introgression (Gay *et al.* 2008). Another category is defined by the presence of intermediate or "flat" genotypic distributions, with an even mixture of parental and hybrid genotypes on the hybrid zone (Jiggins & Mallet 2000). Finally, in bimodal hybrid zones the distribution of genotypes is represented by a higher number of pure individuals with few intermediates, which suggests a restriction to gene flow with strong prezygotic barriers (Jiggins & Mallet 2000). Considering the biological species concept, speciation exists when different species remain distinct in sympatric zones, which is the case of bimodal hybrid zones where the speciation of parental forms is almost complete.

Hybrid zones have been widely used in the study of evolutionary processes (Hewitt 1988a; Harrison 1993; Dowling & Secor 1997), as they can provide a present view of the process of divergence and speciation. For this reason, they have been used for systematic inference and to understand different modes of speciation (Jiggins & Mallet 2000). Indeed, their study can enhance our knowledge about the evolutionary mechanisms of reproductive isolation, selective forces against hybrids and the genetic basis of local adaptation (Kawakami & Butlin 2001). Also, hybrid zones allow the study of several other biological characteristics without long laboratory crossing as they already present many interactions between characters and highly recombined genotypes in a natural context. Examples are the distribution of rare alleles, morphological asymmetry, gene expression, testicular dysfunction, mate preferences, dispersal rate, population size and patchiness, competition, and the effects of parasites, to name just a few (see Hewitt 1988 and references therein).

Prezygotic barriers seem to be important in the formation and maintenance of hybrid zones (Coyne & Orr 2004), especially behavioural differences that cause assortative mating which can reduce gene flow between hybridising taxa. This type of barriers has been documented in several hybrid zones (e.g. between the flycatchers *Ficedula hypoleuca* and *F. albicollis* (Sætre *et al.* 1999)). Also, habitat segregation and temporal isolation can pose strong barriers to gene flow (Kawakami & Butlin 2001). These barriers can exist in combination with postzygotic barriers, causing reduced fitness on hybrids, which can result in strong linkage disequilibria and a deficiency in heterozygotes near the centre of the hybrid zone. However, the contribution of each barrier to the maintenance of the proportion of hybrids in a stable hybrid zone is normally unknown.

1.3. Using genetic traits to study hybridization

The recognition of the importance of hybridization in nature is largely a consequence of new insights provided by molecular studies. Molecular analyses allow a more rigorous identification of hybrids (Mallet 2005) and an evaluation of the type of hybrids (F1, F2, etc.) (Anderson & Thompson 2002). Moreover, using a large number of genetic markers may be useful in the detection of genomic regions with reduced gene flow, which are expected to contain genes responsible for reproductive isolation and local adaptation (Kawakami & Butlin 2001; Baack & Rieseberg 2007; Gompert *et al.* 2012; Kingston *et al.* 2012). However, it is important to understand the differences between different types of molecular markers, as these may have different modes of evolution, varying in the level of information and in possible subsequent interpretations.

In the past, the analysis of DNA variation basically relied on mitochondrial DNA (mtDNA), as this marker is more abundant than nuclear DNA and easier to use with several techniques (e.g. RFLP, sequencing) (Zhang & Hewitt 1996). In general, mitochondrial DNA is maternally inherited presenting no recombination (Avise 1994; Zhang & Hewitt 1996). Therefore, the analysis of this marker, exclusively, in hybridization studies is not very useful, particularly in cases where the forms involved are cryptic. However, when analysed together with more loci, it may provide particularly interesting information on the dynamics of hybridization and introgression, and hence it has been used in several studies of hybrid zones (e.g. Sequeira *et al.* 2005; Barilani *et al.* 2005, 2007; Pinho *et al.* 2009; Fong & Chen 2010). For example, mtDNA may

provide insights with respect to the types of crossings responsible for gene flow by allowing the identification of the maternal lineage of hybrid individuals (Mallet 2005).

The identification of hybrid individuals and therefore the study of hybridization can only be performed using biparental molecular markers, like microsatellite loci (also known as simple sequence repeats or short tandem repeats). These markers are tandem repeats of one to six nucleotides, repeated a certain number of times, with a characteristic mutational behaviour (Ellegren 2004), and they are typically highly polymorphic due to their high mutational rates (Kelkar *et al.* 2010). These markers present some advantages when compared to other biparental molecular markers (such as single nucleotide polymorphisms, SNP). For example, the accuracy of the genotypes is easy to assess (Pálsson *et al.* 1999); they present a high allelic diversity (Kelkar *et al.* 2010; Guichoux *et al.* 2011); and are more or less easy to use in closely related species (Guichoux *et al.* 2011). On the other hand, one of the main disadvantages of microsatellites is the limited reproducibility among laboratories using different equipment or reagents (Guichoux *et al.* 2011). Microsatellites are widely used since the appearance of the polymerase chain polymorphism (PCR) in the late 1980s in pedigree analysis (e.g. Dow & Ashley 1998; Knight & Turner 1998), for genetic mapping (e.g. Wu & Tanksley 1993), or genetic structure studies (Estoup *et al.* 1995). Also, and more importantly for the objective of this thesis, they have been extensively used in hybridization analysis as they allow the distinction of parental forms and identification of hybrid individuals (Lu *et al.* 2001).

In order to analyse the multilocus genotypes of microsatellites, clustering methods and assignment test can be used (Manel *et al.* 2005). These methods allow the study of hybridization at the individual level, as it is possible to infer the contribution of various genetic groups in the ancestry of the individual (Cornuet *et al.* 1999; Pritchard *et al.* 2000). Recently, with the use of Bayesian methods, it is possible to determine the number of genetic groups in a system and the proportion of the genome of the individual that belong to each of these groups (Pritchard *et al.* 2000; Corander *et al.* 2003). These methods use highly polymorphic molecular markers (e.g. microsatellites) and usually lead to accurate results even without previous information about the group or diagnostic markers (Pritchard *et al.* 2000; Anderson & Thompson 2002; Vähä & Primmer 2006). Therefore, they allow a robust characterization of hybrid systems.

1.4. How do they look like? - Morphological traits in hybridization studies

Morphological traits were, and still are, used to infer the evolutionary mechanisms underlying phenotypic characteristics in order to understand how diversity appears and investigate what factors determine its distribution across different geographical and temporal scales (Kaliontzopoulou *et al.* 2011). The morphological characteristics of the organisms are essential to the survival and successful reproduction of individuals, as some morphological traits have important known biological functions such as reproductive behaviour or habitat use. As an example, in some lizards, the femoral pores as source of pheromone secretions could be related with territory acquisition, reproductive signalling (Carretero & Llorente 1993) and recognition (Barbosa *et al.* 2005). Similarly, subdigital lamellae can be related with habitat use (Glossip & Losos 1997). Also, the study of morphological traits is important as developmental changes have consequences at the morphological expression and function affecting the fitness of the organism (Klingenberg 2002).

1.4.1. Advantages and limitations of morphological traits

The application of morphological trait studies in hybrid zones could be beneficial in the identification of pre- and post-zygotic barriers. Usually, in hybrid zones, we can expect that hybrids (at least first generation ones) may be morphologically intermediate between the parental individuals. However, even first generation hybrids are not always morphologically intermediate (Allendorf *et al.* 2001) while intermediate individuals are not necessarily hybrids (Wilson 1992). In the past, when molecular tools were still not available, morphological patterns were used to assess hybridization. However, such an approach has probably underestimated the frequency hybridization and backcrosses in the wild (Mallet 2005), as hybrid individuals are now known to exist even when morphologically intermediates are not present, and there is an evident difficulty in the identification of backcrossed individuals based on morphological traits due to their morphological similarities to one of the parental forms. Therefore, the use of morphological traits exclusively is probably not the best approach in the identification of hybrid individuals. Other limitations of these traits are the non-additive genetic variance for phenotypic traits; pleiotropy, which could be wrongly interpreted as linkage disequilibrium; and phenotypic plasticity that can mask the relationship between allelic and phenotypic clines (Gay *et al.* 2008).

Today, the prime advantage of the study of morphological traits in hybrid zones is not the identification of hybrid individuals. In fact, this is not a common practice. Instead, morphological studies in these areas are used to aid in understanding the evolutionary mechanisms involved in morphological evolution during species differentiation and the maintenance of morphological differences. For example, if the proportion of hybrid individuals is low (when compared to parental forms) and no morphologically intermediate individuals are found even in first generation hybrids, we are probably detecting the effect of strong selection against hybrids and/or assortative matting. Also, hybrids could present anomalies which can be a target for selection. A different approach involves the investigation of the similarities between parental species forming the hybrid zone. If parental forms are morphologically more similar in sympatric zones than in allopatric areas this could indicate extensive introgression or adaptation to similar selective pressures (Martínez-Freiría *et al.* 2008). On the contrary, if parental forms are morphologically more dissimilar this could indicate reinforcement which would suggest the existence of prezygotic barriers (i.e. different morphologies can help avoid interspecific mating).

As is the case for genetic traits, a bimodal distribution of phenotypic characters is expected in the presence of strong selection against hybrids and/or assortative mating combined with high dispersal (Jiggins & Mallet 2000; Gay *et al.* 2008). Hybrid zones can also present weaker selection and assortative mating but morphologically intermediate individuals could not be present. On the contrary, when extensive introgression is present, phenotype frequencies tend to a unimodal distribution, showing a hybrid swarm. Phenotype distribution can also be important in the determination of the movement in the hybrid zone, aiding to establish if the zone is stable which is an assumption e.g. of cline analysis (Barton & Hewitt 1985, 1989; Gay *et al.* 2008). The detailed examination of the spatial distribution of phenotypes allows the investigation of selection across hybrid zones. Despite the advantages of study the phenotypic distribution in hybrid zones this type of characters are not regularly analysed.

1.4.2. Developmental instability

Morphological characters can be used in hybrid zones, not only to detect intermediate individuals, but also to assess developmental instability through the study of fluctuating asymmetry (FA). Normally, organisms have the tendency to present developmental stability i.e. a tendency to follow a specific developmental pathway

under a particular set of conditions (Willmore *et al.* 2007). In bilaterally symmetric organisms, the same genome controls the development of the right and left side of the symmetric bilateral features and, as both are developed in the same environment, they are expected to present a perfect symmetry (Palmer & Strobeck 1986). As such, small, random deviations from perfect symmetry (i.e. fluctuating asymmetry) have been used as a measure of developmental instability (Palmer 1994; Palmer & Strobeck 1986). FA has been associated with lower fitness (Palmer & Strobeck 1986) and it can increase due to extrinsic (environmental) or intrinsic (mostly genetic) stress factors (Palmer & Strobeck 1986; Palmer 1994). Developmental instability and increased FA have been detected in a series of species associated to different causes such as pollution (Eeva & Tanhuanpää 2000), parasites (Møller 1992), extreme temperatures (Savage & Hogarth 1999) and hybridization (Wilsey *et al.* 1998).

Therefore, the analysis of FA could be a useful tool for investigating hybrid zones, as an increase in the degree of FA across hybrid zones could be an indication of selection against hybrids. Indeed, hybrid individuals may exhibit higher developmental instability, which may result in negative selection. Despite its potential, fluctuating asymmetry has not been extensively used in hybrid studies, but some examples exist (e.g. *Mus musculus*, where increased developmental stability was detected in populations with admixture (Alibert *et al.* 1994); and *Apis mellifera*, where no significant differences were found in the level of asymmetry (Smith 1997)).

1.5. The spatial structure of hybrid zones

The spatial distribution of genotypes and phenotypes across hybrid zones as well as their coincidence/discordance in the space may allow the determination of the processes generating and maintaining hybrid zones (Barton & Hewitt 1985). The study of the spatial distribution of neutral markers combined with the variation of phenotypic traits would allow the study of which traits are involved in reproductive isolation and the estimation of the levels of selection (Brumfield *et al.* 2001; Gifford 2008; Benson *et al.* 2012).

From a geographical perspective, hybrid zones can be classified as extended areas of overlap, narrow contact zones or mosaic zones (Arnold 1992). Thus, phenotypes and genotypes across a hybrid zone could be spatially structured as clines with a continuous transition from one parental form to the other (Barton & Hewitt 1985). On the other hand, hybrid zones could be spatially structured as a mosaic distribution

of phenotypes/genotypes across the landscape, which is normally related with environmental heterogeneity (Britch *et al.* 2001). The position of hybrid zones could be related with distinct habitat or environmental gradients. However, they may also exist in areas with a uniform environment with no apparent environmental differences (Kawakami & Butlin 2001).

1.5.1. Landscape genetics

Landscaped genetics is defined as the "amalgamation of molecular population genetics and landscape ecology" (Manel *et al.* 2003). Therefore, a landscape genetics approach is applied in all the studies that combine population genetic data with data on landscape structure (Holderegger & Wagner 2006). The combination of improvements in molecular genetics (e.g. Bayesian approaches which allow an individual-based analyses) and new statistical tools (e.g. geostatistics) have promoted the development of this recent field (Manel *et al.* 2003). One of the important characteristics of landscape genetic is scale. It is defined by the species-specific biological and ecological processes under study, and by the spatial dimension at which we are able to sample (Holderegger & Wagner 2006). For example, we can investigate the genetic structure at a fine-scale through an individual-based approach that infers gene flow directly from genetic distances among individuals (Holderegger & Wagner 2006, 2008).

Studies performed in a landscape genetic framework provide information on the spatial configuration of the microevolutionary processes, providing a better inference on the nature of gene flow, genetic drift and selection than classic population genetics methods (Manel *et al.* 2003). This approach also allows the detection of genetic discontinuities and their correlation with landscape environmental features, such as barriers (Manel *et al.* 2003). Therefore, we can use this kind of studies in hybrid zones to enhance the detection of barriers to gene flow related to the environment in the area of contact. At the same time, a spatial approach may allow associating the genetic differences between parental and hybrid individuals with the characteristic of the landscape. With the use of landscape genetics we can investigate the spatial structure of genotypes in hybrid zones through the combination of population genetics analyses (e.g. genetic distances between individuals) with the geographical coordinates of parental and hybrid individuals.

1.5.2. Cline theory

Cline theory provides a powerful conceptual framework in which it is possible to understand the maintenance of hybrid zones and estimate the influence of dispersal and selection (Barton & Hewitt 1985; Gay *et al.* 2008). A cline can be defined as the estimated frequencies of alleles or phenotypes across a spatial gradient (Kruuk *et al.* 1999). When the cline transition of a genetic or phenotypic trait is narrow when compared with the capacity of dispersal of the organism, the influence of that trait to reproductive isolation is stronger as compared to traits characterized by wider clines. Therefore, cline analyses allow the identification of the patterns of introgression and the determination of the factors contributing to reproductive isolation between different species in a hybrid zone (Barton 1979; Rieseberg *et al.* 1999) and they are therefore a powerful tool for examining the influence and nature of reproductive barriers. The shape of a cline can be studied trough the analysis of the sigmoid shape at the centre of the cline, and the two exponential decay curves at both sides of its centre. These parameters allow the estimation of coincidence and concordance of clines at different traits and the investigation of the influence of the selection against hybrids (Szymura & Barton 1986).

The type of selection operating in hybrid zones is important for interpreting cline patterns. Indeed, whether selection is endogenous or exogenous (that is, if it acts upon different alleles in different genetic backgrounds or different environments, respectively (Kruuk *et al.* 1999)). In exogenous selection, fitness is defined based upon the environment and determined by environmental gradient across the area. On the other hand, in endogenous selection fitness is determined by within-genome interactions such as epistasis and it is independent of the environment. When only endogenous selection is present, fitness is independent of the location but it could still be mediated by interactions with the environment such as different abilities to escape predators (Kruuk *et al.* 1999 and references therein). In a tension zone endogenous selection is reflected in strong selection against hybrids being balanced by dispersal of parental genotypes (Barton & Hewitt 1985). However, when exogenous selection is the strongest force it is reflected in differential natural selection across an environmental gradient (Moore *et al.* 1993).

The distinction between the two types of selection is not straightforward as they result in similarly shaped clines. Still, some specific characteristics of the clines could be analysed (Moore *et al.* 1993). When exogenous selection is a strong selective force, the position and width of the cline must be related with the position and width of the

ecotone. On the other hand, if endogenous selection is shaping the contact zone, the cline should be independent of the environment, and different and independent traits should not exhibit the same cline shape. The distinction of these two types of selection is important as it would allow the understanding of the factors responsible for divergence. In theory, it will allow to know if species divergence occur due to the accumulation of genetic incompatibilities or by the diversity of alternative ecological niches (Kruuk *et al.* 1999). However, there is no apparent reason for either type of selection to be exclusive. In fact some studies show the influence of both types of selection simultaneously in hybrid zones (e.g. salamanders (Alexandrino *et al.* 2005) and chickadees (Bronson *et al.* 2003)).

Different forms of clines exist in continuous habitats. For instance, the cline may exhibit a smooth transition from one parental type to the other. In such cases, dispersal is negligible and selection maintains a stable equilibrium at each locality. This type of clines could be related with increased fitness in hybrids (Barton & Hewitt 1985). Other clines exhibit a steep transition across the hybrid zone. This type of clines are maintained by a balance between dispersal and selection and can be related with differences in the environment or selection against hybrid (Barton & Hewitt 1985).

1.6. Wall lizards of the genus *Podarcis*

Wall lizards of the genus *Podarcis* belong to the order Squamata and to the family Lacertidae (Wagler 1830), and they are probably the best studied group of reptiles in the Mediterranean (Camargo *et al.* 2010). Species of the genus are distributed around the Mediterranean basin, being one of the most abundant and widely distributed reptile genera in Europe and North Africa (Arnold 1987). The genus is composed by small sized lizards, very similar in morphology and ecology (Arnold 1973, 1990). *Podarcis* are habitat generalists (Arnold 1987) and most of them use rocks, trunks, vegetation or bare ground for thermoregulation, foraging and shelter.

The systematics of the genus is a matter of discussion until today in terms of relations between species and delimitation of species. Therefore, several studies exist, both focusing on the entire genus (e.g. Harris & Arnold, 1999; Oliverio, Bologna, & Mariottini, 2000; Harris *et al.,* 2005) or on a specific geographical group (e.g. Harris & Sá-Sousa 2001; Harris & Sa-Sousa 2002; Capula & Ceccarelli 2003; Poulakakis & Lymberakis 2003; Pinho *et al.* 2006, 2007a; Lima *et al.* 2009; Kaliontzopoulou *et al.* 2011). Currently, this genus comprises 23 species (Uetz *et al.* 2014), some of them

added recently (e.g. Geniez et al. 2014). Some of the species are critically endangered (Capula *et al.* 2002), while other are successful colonizers (Corti *et al.* 1999).

Iberian and North African wall lizards present a complex evolutionary history which is probably related with their cryptic diversity (Kaliontzopoulou *et al.* 2011). Mitochondrial DNA reveals that *Podarcis* wall lizards distributed across Iberia and North Africa, with the exception of *P. muralis*, form a monophyletic clade (Harris & Arnold 1999; Oliverio *et al.* 2000). Currently, six species are recognized (Kaliontzopoulou *et al.* 2011): *Podarcis bocagei* (Seoane, 1984); *Podarcis carbonelli* (Pérez-Mellado, 1981); *Podarcis vaucheri* (Boulenger, 1905); *Podarcis liolepsis* (Boulenger, 1905); *Podarcis guadarramae* (Boscá, 1916) (former *P. hispanica* type 1); and. *P, virescens*, (former *P. hispanica* type 2) Geniez *et al.* 2014. Normally, the lineages observed in mtDNA are also recovered by nuclear markers (Pinho *et al.* 2007a).

As mentioned above, species of the genus *Podarcis* are habitat generalists (Arnold 1987). However, in the Iberian and North African clade some species show some degree of specialization, as is the case of *P. guadarramae* (Sá-Sousa *et al.* 2002; Kaliontzopoulou *et al.* 2012a) which is very saxicolous. In terms of distribution, the majority of lineages are parapatric, with some exceptions in the pairs *P. bocagei*/*P. guadarramae guadarramae* and *P. carbonelli*/*P. virescens*, which are sympatric or even syntopic. Sympatry is rare but also present between *P. carbonelli* and *P. virescens* in the Iberian Central System and between *P. vaucheri* and *P. carbonelli* in Doñana National Park (Carretero 2008).

In terms of morphological characters, the species of this genus are highly variable in size, shape, scalation and colour patterns (Arnold *et al.* 1978; Harris & Sá-Sousa 2001; Sá-Sousa *et al.* 2002; Kaliontzopoulou *et al.* 2012b). Moreover, this variability is not only present between species but also between populations (Kaliontzopoulou *et al.* 2012b). Finally, sexual dimorphism is strong in the genus (Kaliontzopoulou *et al.* 2007, 2008).

1.7. Hybridization between *Podarcis* species

Throughout the investigation of the systematics of the genus, it has been repeatedly suggested that forms of *Podarcis* are not completely reproductively isolated, but the extent and dynamics of gene flow between them is far away from being totally understood. Overall, hybridization between different forms of *Podarcis* exist but is

limited. In nature, hybrid individuals were identified in several pairs: *Podarcis wagleriana/Podarcis sicula* (Capula 1993); *P. sicula/P. tiliguerta* (Capula 2002); *P.* sicula/P*. raffonei* (Capula *et al.* 2002); *P. bocagei/P. guadarramae lusitanica* (Arntzen & Sá-Sousa 2007; Pinho *et al.* 2007a); *P. vaucheri/P. guadarramae* (Pinho *et al.* 2008); and between *P. bocagei and P. carbonelli* (Pinho *et al.* 2009). The later pair of species has also been shown to hybridize in captivity (Galán 2002). In terms of morphology, when hybrids or introgression where detected, intermediate individuals are rarely found (Pinho 2007; Pinho *et al.* 2009). However, when morphological intermediate individuals where detected between *P. sicula* and *P. tiliguerta*, molecular analysis confirmed a hybrid origin (Capula 2002).

The detection of hybrid individuals between different species of *Podarcis* suggests that the species of the genus may not be completely reproductively isolated (Pinho *et al.* 2009), making them an interesting system for the study of the nature and effectiveness reproductive barriers.

1.8. Study case: *P. bocagei* and *P. carbonelli*

Podarcis bocagei and *P. carbonelli* are two species endemic of the western Iberian Peninsula that were used as model species in the present study.

P. bocagei is a medium sized (mean SVL: males – 56.9mm; females – 54.7mm) insectivorous lizard (Galán 2009), endemic to the North-West of the Iberian Peninsula (Pinho 2008 in Loureiro *et al.* 2008). Males present green dorsal patterns and brown sides with a normally yellow/orange colour and females have a brownish colour all over the body (Galán 2009). The species exhibits sexual dimorphism in some characters such as larger heads in relation to body size in males (Kaliontzopoulou *et al.* 2007). Usually, females of this species can lay from 2 to 7 eggs in 1-3 clutches (Galán 1999), and larger females tend to start breeding earlier in the season and to lay more eggs and more clutches (Galan 1997). *P. bocagei* generally occupies humid habitats, being present in Atlantic areas, where it can occupy a variety of habitats such as dune areas, anthropogenic habitat or even forests (Galán 2009; Kaliontzopoulou *et al.* 2010).

In terms of phylogeography, *P. bocagei* shows little geographical substructure, which seems to indicate extreme glacial retractions and postglacial northern expansions in multiple glacial refugia (Pinho *et al.* 2011). Historically, *P. bocagei* seems to have undergone an extremely rapid demographic growth from a single source in a refugial region smaller than its actual distribution. Therefore, the present range of the

species must have been colonized in the past 10000 years, during the most recent period of climate warming and aridification (Pinho *et al.* 2007b).

P. carbonelli was initially described as a subspecies of *P. bocagei*. However, molecular analyses reveal high levels of differentiation, which supported its elevation to the species level (Harris & Sá-Sousa 2001). This species has a smaller size than *P. bocagei* (mean SVL: males – 50.7mm; females – 49.4mm;Sá-Sousa 2009). Males have a green dorsal pattern and females are totally brownish (Sá-Sousa & Harris 2002). Clutch size ranges between 1 and 5 eggs, with a mean of 2.3 eggs (Pérez-Mellado 1982). *P. carbonelli is* distributed in Western Iberian Peninsula South of the Douro river with a fragmented distribution (Sá-Sousa 2008 in Loureiro *et al.* 2008; Sillero & Carretero 2013). In terms of habitat this species occupies the same humid environments as *P. bocagei*. (Sá-Sousa 2001). Evolutionary studies on *P. carbonelli* (Pinho *et al.* 2007b, 2011) suggest that the current distribution of the species occupies a smaller area than that in the past, most probably due to climatic modification during Pleistocene and Holocene (Sá-Sousa 2001), and that is still decreasing nowadays (Sá-Sousa *et al.* 2009). This, combined with the increased fragmentation of its distribution, has led to the classification of this species as endangered (Sá-Sousa *et al.* 2009). Currently, the species is more abundant at the Northern part of its distribution, being rare and more locally distributed at the south (Sá-Sousa 2001).

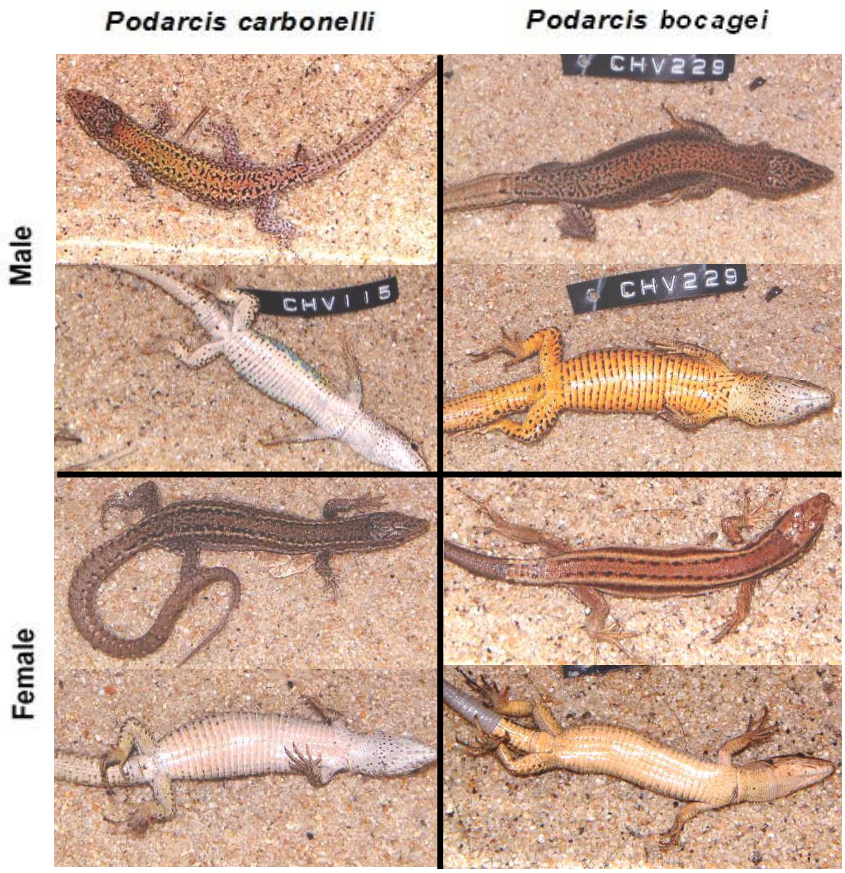


Figure 1 – Dorsal and ventral view of both genders of *P. bocagei* and *P. carbonelli*. Pictures' author – Miguel Angel Carretero.

Regarding their phylogenetic relationships, *P. carbonelli* and *P. bocagei* are closely-related, yet not sister species. According to evidence based on mtDNA, they shared their most recent common ancestor at about 6.5 MYA (Kaliontzopoulou *et al.* 2011). In general, both species are quite similar, both being ground-dwelling lizards. However, some differences do exist, as *P. bocagei* predominates in areas with Atlantic climate and *P. carbonelli* is more constrained to Atlantic or Sub-Atlantic areas that are normally surrounded by large Mediterranean areas (Sá-Sousa 2001). Both species reproduce at the same time of the year, but they appear to have some minor differences in some characteristics such as sperm competition (Carretero *et al.* 2006). Further, some evidences indicate that the two species also differ in their reproductive behaviour, as suggested by the presence of assortative mating (Barbosa *et al.* 2005). In terms of morphology, both species are relatively similar, but they also exhibit significant differences (Harris & Sá-Sousa 2001; Kaliontzopoulou *et al.* 2005, 2012) that remain in sympatry (Kaliontzopoulou 2004) but with less distinction (Pinho *et al.* 2009).

P. bocagei and *P. carbonelli* present mostly a parapatric distribution (Sá-Sousa 2001), being found in strict syntopy in a small area along the north-west coast of Portugal (Carretero *et al.* 2002) in a dune area around the city of Espinho. The primary dune is covered by degraded gramineae (*Otanto maritimi* and *Ammophileton australis*) separated from the secondary dune by an interdune corridor with ephemeral therofytes. The secondary dune is composed by endemic species of the Northwestern Iberian Peninsula with perennial herbaceous and small shrubs (*Iberidetum procumbentis*) separated by clearings composed of small annual plants (*Evax pygmaea*). The vegetation present a coast-inland gradient. The flora distribution is very affected by anthropogenic disturbance and the regression of the coast line. The effect of both factors can be observed in the lack of primary dunes and the substitution of the typical vegetation of the secondary dunes by nitrophilic communities due to human stepping and nitrification of the sand (Carretero *et al.* 2002). At the same time, the expansion of invasive species such as *Carpobrotus edulis* has a negative impact in the vegetation of the area (Carretero *et al.* 2002).

A genetic analysis of the sympatric area between *P. bocagei* and *P. carbonelli* allowed the detection of hybrids with some level of introgression (Pinho *et al.* 2009). Despite the presence of hybridization, a high proportion of parental forms were detected, which suggested the existence of a bimodal hybrid zone with strong barriers to gene flow. Also, a preliminary morphological analysis was performed but no evidence for the existence of intermediate phenotypes in admixture individuals was found. The study put a first basis, but also left several question about the influence and

nature of reproductive barriers maintaining species integrity open, indicating the necessity of a multidisciplinary study of the hybrid zone.

1.9. Objectives

The main aim of the present study is to elucidate the ecological and evolutionary mechanisms promoting reproductive isolation by performing a microscale analysis of the hybrid zone between *Podarcis bocagei* and *Podarcis carbonelli* using a multidisciplinary approach. Ultimately, we want to understand which forces are shaping the hybrid zone and to infer which type of reproductive barriers shape the composition and spatial distribution of the hybrid zone. As a first step, we want to assess the genetic composition of the hybrid zone through the identification of the hybrid individuals and parental forms. Also, we want to determine the morphological characteristics of hybrids to test two hypotheses: 1) if phenotypes of hybrid individuals are intermediate and if morphological traits follow genetic patterns of admixture, in order to understand the influence of hybridization on the phenotype; 2) if fluctuating asymmetry is higher in hybrids versus parental forms which would be an indication of increased developmental instability. Finally, we want to evaluate the existence of geographical segregation between pure forms and hybrids inferring the magnitude of barriers to gene flow at a microscale by performing a fine-scale spatial analysis of genetic and morphological traits at the individual-level.

2. Genetic and morphological characterization of a contact zone between two species of wall lizards

2.1. Introduction

Speciation can be defined as the development of reproductive isolation between diverging taxa (Turelli *et al.* 2001; Coyne & Orr 2004; Wiens 2004; Butlin *et al.* 2012). Ecological, physiological, life-history and behavioural differentiation between different populations or incipient species can gradually promote the establishment of reproductive isolation, eventually leading to speciation (Coyne & Orr 2004). The genetic basis of reproductive isolation is usually thought to result from epistatic incompatibilities in genes responsible for physiological, anatomical, ecological or behavioural isolation (Turelli *et al.* 2001; Coyne & Orr 2004). Gene flow becomes more difficult as the number of genes that present incompatibilities between diverging taxa increase, due to the co-adaptation of different loci within populations (Orr 1995).

Hybrid zones are defined as "narrow regions where genetically distinct populations meet, mate and produce hybrids" (Barton & Hewitt 1985), and they can be seen as natural laboratories where hypotheses about the mechanisms of reproductive isolation can be tested and put into context, improving the knowledge about the nature and origin of species (Hewitt 1988b; Abbott *et al.* 2013). These zones allow us to investigate the diffusion of genes between diverging taxa, to explore the genetic and reproductive barriers promoting divergence and consequently to help in the identification of the mechanisms responsible for speciation and understand how these mechanisms evolve (Barton & Hewitt 1985; Hewitt 1988b; Mallet 2005; Abbott *et al.* 2013). In fact, we can study reproductive isolation and understand the genetic differences and selective forces that separate and maintain the taxa of interest due to the potentially wide range of genotypes that can be found (Barton & Hewitt 1985). When different populations meet in a contact zone they can be completely reproductively isolated if they cannot produce viable or fertile offspring. On the other hand, they may present a scenario of free admixture and eventually converge into a single population. An intermediate scenario occurs when the populations interbreed in a narrow hybrid zone, maintaining their genetic integrity over most of their distribution area (Barton & Hewitt 1985; Jiggins & Mallet 2000; Abbott *et al.* 2013).

The study of hybrid zones may be favoured from the investigation of the phenotypic characterization of the individuals, and may be important to understand which types of barriers to gene flow are promoting reproductive isolation. Specifically, we can use phenotypic traits to investigate introgression and selective forces acting on hybrid phenotypes, since the expected wider range of morphologically variability caused by admixture may help to uncover specific traits associated with reproductive barriers. Also, the analyses of morphological traits in hybrid zones could provide information about the fitness of the individuals as the presence of abnormalities or increased fluctuating asymmetry (small random deviations from perfect bilateral symmetry) on hybrid individuals could indicate a high developmental instability and, ultimately, lower fitness (Palmer & Strobeck 1986, 2003; Palmer 1994). Therefore, a multidisciplinary approach is necessary to analyse the potential mechanisms promoting reproductive isolation in hybrid zones. Multidisciplinary approaches are also important as a high number of secondary contact zones are being found between cryptic species by phylogeographic studies (Hewitt 2001). These cryptic species can result either from speciation processes that are not followed by morphological differentiation or from our failure to detect the existing differences in morphology (Fritz *et al.* 2006; Barbosa *et al.* 2006). In any case, the study of hybrid zones between cryptic species allows to uncover the existence of pre and post-zygotic barriers increasing the knowledge about species boundaries and speciation processes (Phillips *et al.* 2004; Leaché & Cole 2007; Pinho *et al.* 2009; Stuart-Fox *et al.* 2009).

Iberian and North African wall lizards (*Podarcis)* are a complex of closely related, and overall cryptic species which are still not fully reproductively isolated. In fact, gene flow has been suggested to occur between several pairs of species (e.g. Pinho *et al.* 2008). Therefore, species of this genus seem to be a good model to study the evolution of reproductive isolation.

Although gene flow in contact zones has been suggested to be a frequent phenomenon in this clade, only one hybrid zone between parapatric species - *P. bocagei* and *P. carbonelli* - has been documented and studied. These two species are closely-related, yet not sister taxa. According to evidences based on mtDNA, they shared a recent common ancestor between 6.37 and 8.03 million years ago (Kaliontzopoulou *et al.* 2011). The two species are very similar morphologically but different in size, pholidotic and biometric characters (although with some overlap) (Kaliontzopoulou *et al.* 2008, 2012b) and colouration (Sá-Sousa & Almeida 2000), exhibit similar but not equal ecological requirements (Sá-Sousa 2001), and reproduce at the same time of the year, although with some divergence in life history (Carretero *et al.* 2006; Pinho *et al.* 2009). Due to its morphological similarity, these two wall lizards

were initially classified as a single species (Pérez-Mellado 1981), until their recognition as genetically and morphologically differentiated groups (Sá-Sousa and Harris 2002). Despite their morphological similarity, the two species can be distinguished with relative confidence using the right framework and combination of morphological traits (Kaliontzopoulou *et al.* 2005, 2012b).

These two wall lizards share a limited region of their distribution range, as *P. bocagei* is endemic to the North-west of the Iberian Peninsula (Carretero *et al.* 2002; Pinho 2008) and *P. carbonelli* has a fragmented distribution in Western Iberian Peninsula, South of the Douro river (Sillero & Carretero 2013). Indeed, they are only known to overlap in the coastal sandy area of Espinho (Carretero *et al.* 2002) (Figure 2). A previous study (Pinho *et al.* 2009) demonstrated the existence of hybridization between the two species in this contact zone, with residual evidence for introgression in nearby localities. The study of this hybrid zone revealed a highly bimodal genotype distribution, suggesting the existence of strong barriers to gene flow. However, the low sample size used in this study, its sampling scheme (artificially balanced between the two species irrespectively of their frequency in the study area) and the low frequency of hybrids detected limited its ability to document admixture patterns or hybrid properties which could shed light on isolating mechanisms between the two species. A morphological characterization of hybrids was also performed but was affected by the poor resolution in overall species distinction.

Accordingly, in the present study we address the problems from previous analyses by performing an intensive and careful sampling of this area. Not only we aimed at obtaining a much higher sample size, but we sampled proportionally to the relative abundance of each species and their hybrids in the study area, as well as of each sex or age cohort. Both of these features allow more statistically supported inference regarding the dynamics of this hybrid zone. We have analysed these samples using a multidisciplinary approach that combines a genetic and morphological perspective to obtain insights on the mechanisms limiting gene flow and maintaining reproductive isolation between the two species. Specifically, we want to assess whether the general patterns regarding the genetic composition of the contact zone are consistent with previous results not using such a careful sampling scheme or if previous biases had important consequences in our understanding of the dynamics of this hybrid zone. Moreover, if hybridization is detected, we want to investigate its phenotypic consequences. Specifically, we want to test if hybrids are morphologically intermediate or abnormal in any aspect, and if they present a higher developmental instability than the parental forms. We hope these data will help to understand the

evolutionary dynamics of this interesting contact zone and shed light on the reproductive barriers that are in action.

2.2. Methods

2.2.1. Study area and sampling method

Previous studies have shown that introgression is almost restricted to the northwestern Portuguese coast, around the locality of Espinho (41º01´39.76´´N, 8º38´42.04´´O), the only known area of sympatry between *P. bocagei* and *P. carbonelli* (Figure 2) (Carretero *et al.* 2002, Pinho *et al.* 2009). Our sampling was therefore focused on this locality, covering a stretch of dune, with an extension of approximately 2.3 kilometers long.

The wall lizards of the area present the typical behavior of ground-dwelling lizards, mainly using the scrubland for taking refuge, feeding, and thermoregulating, the open sand areas to disperse and wooden passes probably for dispersal and refuge. *P. bocagei* shows a tendency to occupy areas with higher vegetation, while *P. carbonelli* uses more clear and open areas with lower vegetation (Carretero *et al.* 2002). However, both species overlap extensively and use very similar resources.

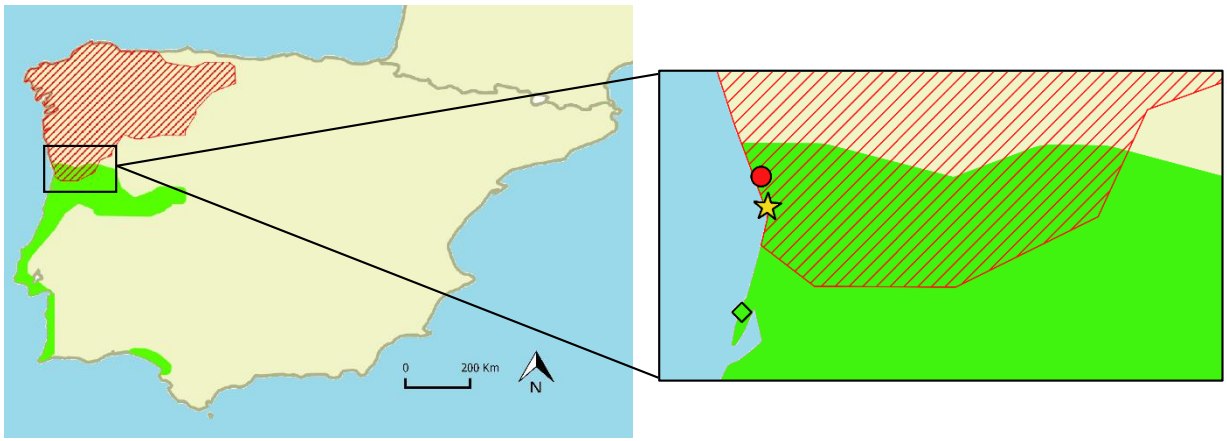


Figure 2 - Distribution of *Podarcis bocagei* (red) and *Podarcis carbonelli* (green) in the Iberian Peninsula and contact zone between them (yellow star). The area of overlap between both species is indicated in a large scale by the distribution maps of both species. However, this large overlap area is not exact due to the use of 10x10Km squares. Both species are only found in sympatry in a small area in Espinho. The sampling locations Madalena (red dot) and Torreira (green rhombus) are also identified.

Individuals were captured along and around the dune area using a noose (García-Muñoz & Sillero 2010) or by hand. The data and precise geographic coordinates (WGS1984 datum) of each captured individual were recorded with a high precision global position system GPS (Garmin E-trex) with an error of approximately 4 meters. For each individual we recorded snout-vent length (SVL), as a measure of total body size and for its classification as adult or juvenile. Several high-resolution photographs of the individuals were taken (dorsal, ventral and both laterals of the entire animal as well as of the head; chest; cloacal area; and the fourth toe on both sides of the body) using a digital camera (Canon PoweShot A495). Finally, a small tip of the tail of each specimen was collected and preserved in 96% ethanol in order to perform DNA analyses. The individuals were released at the exact point where they were captured.

Sampling was performed during 2013 in two different seasons: in spring (from April to June), during the reproductive season, and in September when dispersal of juveniles takes place and newborns can be easily collected.

Our sampling scheme aimed at collecting as many individuals as possible, without performing any selection of individuals with respect to species, size, location or sex. A total of 202 individuals were sampled in the contact zone including 165 adults and 37 juveniles. The individuals were considered adults if their size was larger than that of the smallest reproductively active individual (Carretero *et al.* 2006; Galán 2009; Sá-Sousa 2009). A provisory assignment of individuals to each species was done based on various empirical diagnostic characters which include size, colour pattern and head shape (Pinho *et al.* 2009) either during sampling or afterwards, by photography observation. According to this visual identification, our final sample included 141 *P. carbonelli*, 57 *P. bocagei* and 4 non-identified individuals of which only tails were collected. The identification of the sex of each individual was based on colour patterns, head shape and presence of developed femoral pores in males (Ferrand de Almeida *et al.* 2001). Both sexes were sampled with 74 adult females, 87 adult males, 17 juvenile females, 19 juvenile males. The sex of 5 individuals could not be inferred due to the difficulty in the identification of the sex in juveniles. Samples from other localities (Torreira (N=25) and Madalena (N=25)), supposedly composed by only pure individuals, were used as a reference.

In order to evaluate the presence of recaptured individuals in our sample we used the software Interactive Individual Identification System (I3S CLASSIC 2.0, Van Tienhoven *et al.* 2007) which performs photographic identification. This technique is based on the identification of regular and individually specific patterns of ornamentation in well-identified body regions. We used the pattern of the intersections among pectoral scales because the points of intersection among scales are unique to each individual

and highly divergent between individuals in number and position, working as fingerprinting in lizards (Sacchi *et al.* 2010). This method is a good alternative to toe clipping as it is non-invasive and less likely to be affected by misclassification. Using this method we found four recaptured individuals (from the 202 individuals sampled) in our sample that were excluded.

2.2.2. DNA extraction

The tail tips of sampled individuals were used to extract genomic DNA using EasySpin® Genomic DNA Tissue Kit (Citomed, Lisbon) and QIAamp® DNA Micro Kit (Qiagen, Hilden) for degraded or small samples, following the manufacturers' protocol. The quality and quantity of extracted DNA were assessed by gel electrophoresis. Extracted DNA bands were visualized in a UV transilluminator (Bio-rad) and diluted with ultra-pure water according to visual quantification, to avoid inhibitors during polymerase chain reaction (PCR). DNA concentration and purity was then assessed using a Thermo Scientific NanoDrop 2000 spectrophotometer.

2.2.3. Mitochondrial DNA amplification and analyses

A portion of the mitochondrial region 12S rRNA was amplified using the primers 12Sa and 12Sb described in Kocher *et al.* (1989), in order to identify the maternal lineage of each individual. This fragment allows the distinction of the mtDNA of both species as they form monophyletic groups with a pairwise difference of approximately 3% (Harris & Sa-Sousa 2002). Amplification was carried out in a T100™ Thermal Cycler (Bio-Rad) with a final volume of 10 µl containing 0.8µM each primer, 5µl Mytaq (Bioline), 2.2µl distilled H2O, and approximately 40µg of genomic DNA. Polymerase Chain Reaction (PCR) conditions consisted in an initial step of 10 min at 95ºC followed by 7 cycles of denaturation at 95ºC for 30sec, touchdown annealing decreasing 0.5ºC every cycle between 53-50ºC for 30sec and extension at 72ºC for 30sec. Next, 33 cycles with denaturation at 95ºC for 30sec, annealing at 50ºC for 30sec and extension at 72ºC for 30sec were performed, ending with a final extension at 72ºC for 5min. PCR products were purified and sequenced in a commercial sequencing facility (Macrogen, The Netherlands). Sequences were analysed using SEQSCAPE® 2.5 (Applied Biosystems) and blasted in GenBank to confirm their identity.

2.2.4. Microsatellites

A battery of 26 published microsatellites (Agostini *et al.* 2013) developed for a related species of the genus *Podarcis* (*P. guadarramae lusitanica,* previously named *Podarcis hispanica* morphotype 1A) were chosen based on the reported absence of null alleles and other genotyping errors. The loci were combined into six multiplex sets (Table 1). Allele scoring in microsatellites is based on different sizes, so loci were combined taking into account possible overlaps in size range between markers. Moreover, it is also important to avoid, or at least minimize, primer-primer interaction in a multiplex. Thus, AUTO-DIMER software (Vallone & Butler 2004), which attributes a score that represents the degree of interaction between primer oligonucleotides (higher scores represent higher complementarity) for each primer pair combination, was used, and primer pairs with scores equal to or higher than 7 (recommended value) were separated in different multiplexes. A fluorescent dye (6-FAM, VIC, NED or PET) was added on the 5´ end of the forward primer (PF) of each marker, according to (Agostini *et al.* 2013).

From the 26 microsatellites, four failed to amplify (Ph23_3, Ph37, Ph46_8 and Ph124), so the six multiplexes were done with the remaining 22 markers (Table 1). Multiplex reactions were optimized for primer concentration and PCR conditions using samples of both species. PCR was performed in a T100™ Thermal Cycler (Bio-Rad) with a final volume of 10µl using 5µl of Qiagen Multiplex PCR kit, 1µl of a multiplex mix including all primers of each multiplex (see Table S1 in supplementary material for concentration details) and approximately 50ng of genomic DNA. In order to detect DNA contaminations a negative control was used.

For most multiplexes (Multiplex A, B, C, D and E – Table 1) PCR conditions consisted of an initial step of 15min at 95ºC followed by 15 cycles of denaturation at 95ºC for 30sec, touchdown annealing decreasing 0.5ºC every cycle between 62-55ºC for 1.30min and extension at 72ºC for 45sec. Next, 22 cycles with denaturation at 95ºC for 30sec, annealing at 55ºC for 45sec and extension at 72ºC for 45sec were performed. Following, 8 cycles consisting of a denaturation step at 95ºC for 30sec, annealing at 53ºC for 30sec, extension at 72ºC for 45sec were conducted, ending with a final extension at 60ºC for 30min. For the Multiplex F, PCR conditions were similar except for the touchdown cycles (11 and not 15) in which annealing temperatures decreased 0.5ºC every cycle between 62-57ºC.

After testing for amplification success, PCR products were size-separated on an ABI 3130 capillary sequencer (Applied Biosystems), using the Genescan-500 LIZ (ABI)

as size standard. Alleles were scored using GENEMAPPER 4.0 (Applied Biosystems) and checked manually.

Table 1 - Composition of the six multiplex sets and characteristics of the markers analysed in this study. The allele range corresponds to that reported in Molecular Ecology Resources Primer Development Consortium (Agostini *et al.* 2013) for *P. guadarramae lusitanica*

Locus	Repeat motif	Allele range (bp)	Multiplex
Ph14	$(ACAT)_n$	138-182	MpA
Ph31	$(GATA)_n$	138-192	MpA
Ph32	$(ATTG)_n$	89-137	MpA
Ph59	$(TTC)_n$	110-113	MpA
Ph22	$(TCT)_n$	201-241	MpB
Ph35	$(TACA)_n$	145-175	MpB
Ph39	$(ATCT)_n$	108-144	MpB
Ph170	$(TTC)_n$	210-216	MpB
Ph17	$(TATC)_n$	149-225	MpC
Ph21	$(AGAT)_n$	127-189	MpC
Ph30	(TCTA)n	108-148	MpC
Ph38	$(GATT)_n$	100-136	MpC
Ph50	$(ATGC)_n$	266-282	MpC
Ph25	$(TCTA)_n$	174-200	MpD
Ph83_11	$(ACA)_n$	262-298	MpD
Ph142_8	$(TCC)_n$	269-290	MpD
Ph43	$(AGGG)_n$	136-153	MpE
Ph70	(CTT)n	162-195	MpE
Ph81	$(TGT)_n$	288-321	MpE
Ph128	$(GTT)_n$	218-233	MpE
Ph61	(AAC)n	154-199	MpF
Ph62	$(GTT)_n$	202-235	MpF

In order to estimate the rate of genotyping errors such as allele dropout and false allele scoring, an independent PCR and allele scoring was performed in 23 samples (around 12% of the samples of Espinho).

2.2.5. Microsatellite data analyses

Samples with more than 20% of missing data were excluded from the dataset in order to avoid possible biases. The locus Ph59 was also excluded due to a high rate of missing data (approximately 20%) and because we were not able to amplify any of the samples of the reference population of *P. bocagei* (Madalena), making it impossible to perform assessments of genotyping errors in in tests involving reference populations

(see below). Marker Ph170 was excluded from subsequent genetic analyses because it was monomorphic in the entire sample.

Mean allele dropout and false allele error rates were calculated by PEDANT 1.0 software (Johnson & Haydon 2007), which calculates error rates at each locus by comparing genotypes from two replicate sample batches. Due to the expected effects of hybridization on Hardy-Weinberg and linkage equilibria (HWE and LE, respectively) and to avoid circularity in the definition of test populations, the tests for genotyping errors that involve testing for HWE were performed using the reference samples (Madalena and Torreira). These were performed using MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2004), which was used to identify genotyping errors due to null alleles, short allele dominance (large allele dropout) and stuttering, and the GENEPOP 4.2.2 (Raymond & Rousset 1995; Rousset 2008) probability tests for deviations from HWE and LE in each reference population. All markers with significant evidence for genotyping errors in at least one of the reference populations (after Bonferroni correction, see Results) were excluded from subsequent genetic analyses.

To investigate the power of the microsatellite markers for individual identification, the probability of identity (the probability that two individuals in the population share the same genotype by chance) and the probability of identity in siblings (probability that two (randomly chosen) individuals within a given population have the same genotype on a set of markers) of each loci separately and overall across loci was calculated taking into account sample size corrections using the GIMLET software, version 1.3.3 (Valière 2002).

To identify the admixed individuals several analyses of individual assignment based on multilocus genotypes were performed. First, the software GENETIX 4.2 (Belkhir *et al.* 1996) was used to perform a Factorial Correspondence Analyses (FCA). Next, we used STRUCTURE 2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007; Hubisz *et al.* 2009), which implements a Bayesian model-based clustering algorithm that identifies clusters of individuals that minimize deviations to HWE and LE and estimates, for each individual, the proportion of the genome originating in each cluster. We performed a variety of exploratory analyses, both including and excluding reference populations, testing different numbers of populations (K; from 2 to 10) and with different parameter settings (both correlated and independent allele frequencies) to evaluate the robustness of the results. Final analyses were conducted with 10 independent replicates of the Markov Chain of Monte Carlo (MCMC), each with 10^6 iterations including a burn-in period of 250000 steps.

Additionally, we investigated patterns of hybridization with greater detail using the software NEWHYBRIDS 1.1 (Anderson & Thompson 2002). This software uses a

model-based Bayesian approach to calculate the posterior probability (PP) of multilocus individual genotypes belonging to different admixture categories. We tested the six default categories, corresponding to pure of each parental type, F1 or F2 hybrids, and backcrosses resulting from the mating of an F1 to each of the parental forms, excluding reference individuals. Three runs were performed with 500000 MCMC iterations each and discarding 100 000 iterations as a burn-in.

Simulated datasets were used to evaluate the power of the markers in the assignment of individuals by both clustering algorithms and define appropriate threshold values to assist in decisions regarding the admixed nature of each individual. Simulations were performed using HYBRIDLAB 1.0 (Nielsen *et al.* 2006). For these analyses, individuals assigned unambiguously (individuals with posterior probabilities equal or higher than 0.99) as pure of each parental species in the previous steps were used to simulate the "pure" genotypes. To avoid circularity, we used the results of STRUCTURE for the simulated datasets run in NEWHYBRIDS and vice-versa. The simulated dataset includes the same proportion of classes and total number of individuals of the original dataset were maintained in order to account for the sample size so that simulation conditions mimicked, as much as possible, those applied for the real data set. The simulated datasets were run in NEWHYBRIDS and STRUCTURE under the same conditions as the previous analyses.

Finally, genetic diversity was evaluated in *P. carbonelli* and *P. bocagei* from the contact zone, excluding hybrid individuals (identified using STRUCTURE). Allele frequencies for each locus, as well as observed (Ho) and expected (He) heterozigosities, mean number of alleles (Na), and private alleles (Na private) per locus were calculated using ARLEQUIN 3.5 (Excoffier *et al.* 2005; Excoffier & Lischer 2010). An Analyses of Molecular Variance (AMOVA, Excoffier et al., 1992) was performed using the same software in order to assess population differentiation between both species at the contact zone. For this purpose, a Fst-like calculation (Cockerham & Weir 1984) was executed based on allele frequencies and disregarding allele size.

2.2.6. Pholidotic patterns

The individuals sampled in the contact zone and used in the previous genetic analyses were also analyzed for different morphological traits in order to examine possible particularities of the morphology of hybrid individuals.

A total of 13 characters (Figure 3) used in previous taxonomic studies in lacertids were examined (Kaliontzopoulou *et al.* 2005, 2012b). Continuous pholidotic characters included: collaria (CSN); femoral pores (FPN); gularia (GSN); supraciliary

granules (SCGN); subdigital lamellae under the forth toe (SDLN); supratemporal scales (STSN); number of transversal rows of ventral scales (VSN); supraciliary scales (SCSN); and supralabial scales (SLSN). Also, four categorical characters were recorded: IN_F, contact between the internasal and frontal scales (0 – no; 1 – yes); O_IP, contact between the occipital and interparietal (0 – no; 1 – yes); MASS, presence of the masseteric scale (0 – absent; 1 – present); TYMP, presence of the tympanic scale (0 – absent; 1 – present). All bilateral characters were measured at the right side of the body. The categorical characters did not present enough variation in our sample so they were not used in posterior analyses (see results).

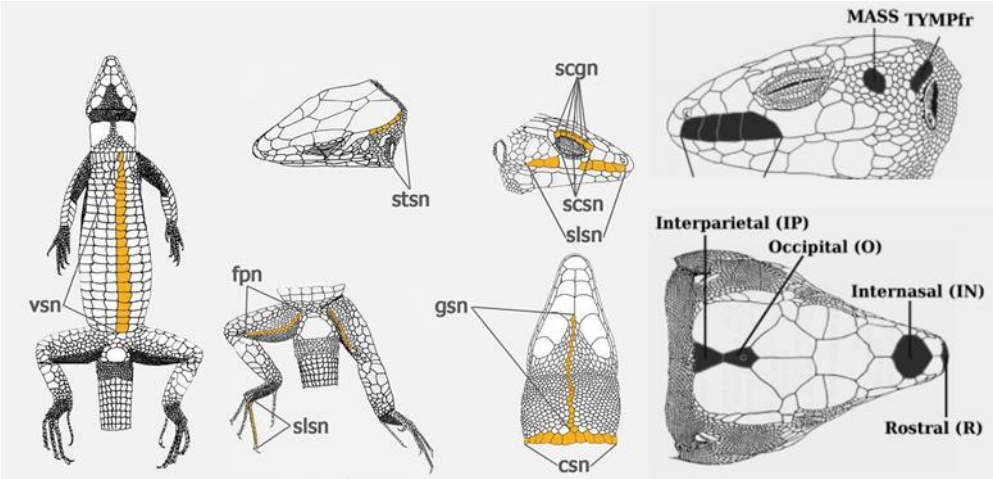


Figure 3 - Ordinal pholidotic characters (orange) and categorical characters (black) recorded. See material and methods for variable abbreviations. Adapted from Kaliontzopoulou *et al*. (2012)

Ordinal pholidotic characters were examined in order to test whether they could be statistically treated as continuous for statistical analyses. Characters were treated as continuous when they presented five or more categories. A descriptive analysis was performed to calculate the means and standard deviation for each character taking into account species (based on the assignments of STRUCTURE) and sex (based on several characters mentioned in methods) in order to describe the variation at each character.

To examine patterns of morphological variation in continuous pholidotic characters it is important to test for the presence of sexual dimorphism as it was already indicated as an important source of variation in these species (Kaliontzopoulou *et al.* 2005). If present, sexual dimorphism can mislead result interpretation due to increased variation in our analysis, for all the studied characters. Also, we need to determine which characters are useful for species discrimination. For that purpose, we first conducted a Factorial Analysis of Variance (ANOVA) of each variable with sex and species (determined by previous genetic analyses and excluding all hybrid individuals)

as factors, as well as the interaction term. As ANOVA results indicated the presence of sexual dimorphism in some characters (see results) and since sexual dimorphism was not the main focus of this study, residual values corrected for sex were calculated and used in posterior analyses.

In order to investigate the main sources of variation in our sample a Principal Component Analysis (PCA) was performed in continuous pholidotic traits. The same analysis was performed, taking into account sex and species, separately. As the aim of these analyses is to construct a discriminant analysis (DA) that maximize the difference between the parental species, only ordinal characters that exhibited significant differences between species were used in posterior analyses. The remaining characters were excluded to reduce variability (background noise) and to increase the discriminatory power between species. A PCA was performed using only the characters that presented significant differences between species in order to visualize variation between individuals (see Figure 6). All previous analyses were performed using the software STATISTICA 10 (StatSoft 2011).

In order to classify potential hybrid individuals pinpointed by genetic analyses, we construct a DA on continuous pholidotic characters using STATISTICA 10 that allows the study of the discrimination between both species and detection of the characters that are most differentiated between them. In order to perform the DA we use allopatric localities of both species (Madalena for *P. bocagei*, N=36 and Torreira for *P. carbonelli,* N=36*,* the same used as reference in genetic analyses, which have similar habitats and are relatively close to the hybrid zone). First, we ran an ANOVA using the factors population and sex in both species separately, in order to test for differences between the allopatric and sympatric populations of the same species and to avoid increasing the variability within each species that could difficult the construction of the discriminant analysis. Subsequently, in order to construct the discriminant function we used the 36 individuals of each allopatric locality and 29 (the number of individuals of the species less sampled) from each species at the contact zone. Because the sample size for *P. carbonelli* was much higher, we used a random sample of 29 *P. carbonelli* individuals to avoid biasing the discriminant function.

In order to investigate the morphological characteristics of hybrids, the posterior probabilities (PP) of every hybrid to belong to each one of the parental species were calculated using the morphology based discriminant function constructed with the pure individuals from allopatric and sympatric areas. Then, the PP assigned to each of the hybrid zone individuals were compared with the proportion of the genome originating from each species, estimated from genetic data using STRUCTURE. In order to obtain a measure of the correlation between the morphological (using DA) and genetic based

(STRUCTURE results) PP at each species we perform a Spearman´s rank correlation coefficient using STATISTICA 10. To perform the analysis individuals were classified using the mtDNA results to classify individuals as *P. bocagei* or *P. carbonelli* (as it provides an independent criterion for group assignment).

2.2.7. Asymmetry analyses

In order to investigate the existence of developmental instability we evaluated fluctuating asymmetry. In these analyses three meristic pholidotic characters were measured in both sides of the body of all individuals from Espinho: femoral pores (FPN); subdigital lamellae on the fourth toe of the hindlimb (SDLN); and supraciliary granules (SCGN). These three traits were chosen because they are easy and fast to quantify and exhibit extensive variability in *Podarcis* populations and species (Kaliontzopoulou *et al.* 2012b) and have been previously used successfully to recover differences in asymmetry in *Podarcis* populations (Lazić *et al*. 2013). In order to account for measurement error all the characters of each individual were measured twice, allowing several days of rest between both counts and randomizing the order of the individuals examined.

In order to test for the presence of directional or fluctuating asymmetry (FA), a two-way ANOVA was performed while accounting for measurement error. The main effects examined were side as a fixed factor and individual as random factor, as well as their interaction. In this ANOVA design, a significant effect of individual represents the variation among individuals. A significant effect of side represents the variation between the right and left side of the body and it indicates the existence of directional asymmetry. Finally, a significant interaction between individual and side indicates the presence of FA, being this the variability of right-left differences amongst individuals (Palmer & Strobeck 1986; Palmer 1994). Analyses for studying asymmetry patterns were performed within each of the three groups (*P. carbonelli*, *P. bocagei* and hybrids, as identified by STRUCTURE) and for each character separately.

Two different asymmetry indexes (AI) were calculated for each of the examined traits in each individual: the first corresponds to the difference between the trait value on the right side of the body, and that on the left side (AI1 = R - L); the second was calculated as the mean of the absolute value of the difference between the right and the left side (AI2 = mean |R – L|) (Palmer & Strobeck 1986; Palmer 1994).

In order to exclude existence of directional antisymmetry (which can inflate the values of the AI), deviations from normality in the AI were tested within each group (two parental forms and hybrids) in all the traits, using a Kolmogorov-Smirnov test.

Dependence of the degree of asymmetry on total body size was tested using a linear regression of each of the two AI values on SVL. Also, to test for dependence of trait size, a linear regression of AIs on mean trait size at each sample (R+L)/2 was performed.

Additionally, we calculated an individual asymmetry index for each trait as the unsigned R - L difference between the log transformed average of trait values across the two replicate counts at each sample, in order to account for measurement error (|ln(Raverage)-ln(Laverage)|) (Palmer & Strobeck 2003). Finally, in order to investigate the existence of differences in the level of FA an ANOVA was performed. The factors tested were group type (Parental and Hybrids), species (*P. bocagei* and *P. carbonelli*), nested in type, where the hybrids were attributed to one of the species according with the assignments of STRUCTURE (individuals were attributed to the cluster with higher PP value), and sex as factors, being the individual FA index described above used as the response variable. All the interaction effects were also calculated. All statistical analyses were perform using STATISTICA 10.

2.3. Results

2.3.1. Assessment of genotyping errors and deviations to equilibrium

Over all samples the number of alleles ranged from 4 (in Ph128) to 25 (in Ph21) and values of expected heterozygosity ranged from 0.379 in locus Ph43 to 0.905 in locus Ph70. Most loci did not show deviations from Hardy-Weinberg or linkage equilibria in the reference populations. There were few exceptions, identified by MICRO-CHECKER as resulting from genotyping errors: these were loci Ph142_8 (stuttering and null alleles) and Ph17 (null alleles) in the population from Madalena, and loci Ph35 (null alleles) and Ph50 (stuttering) in Torreira. These four loci were eliminated from further analyses. Allele dropout and false alleles rates identified using PEDANT were in general low across loci, and present in a small number of loci with a maximum of 3% false allele rate and 5% dropout both in Ph62 marker (Table S2, in supplementary material). Given our caution in disregarding problematic loci and the low error rates inferred for the remaining markers, the influence of genotyping errors in the estimates of admixture presented below is likely very low. The probability of identity estimated using the individuals from Espinho was of 6.77×10^{-18} which indicates that the combination of the battery of microsatellites selected will performed well in individual assignment.

2.3.2. Genetic composition of the contact zone and genetic identification of hybrids

The FCA scores of each individual sampled in the contact area were graphically represented in a bidimensional plot defined by two principal axes that together explained 11.1% of the total genetic variation. The graphic shows that two clusters are clearly separable, with a high degree of correspondence to the preliminary classification of the individuals into the two species. Some individuals are placed between these two clusters, suggesting an admixed origin (Figure 4).

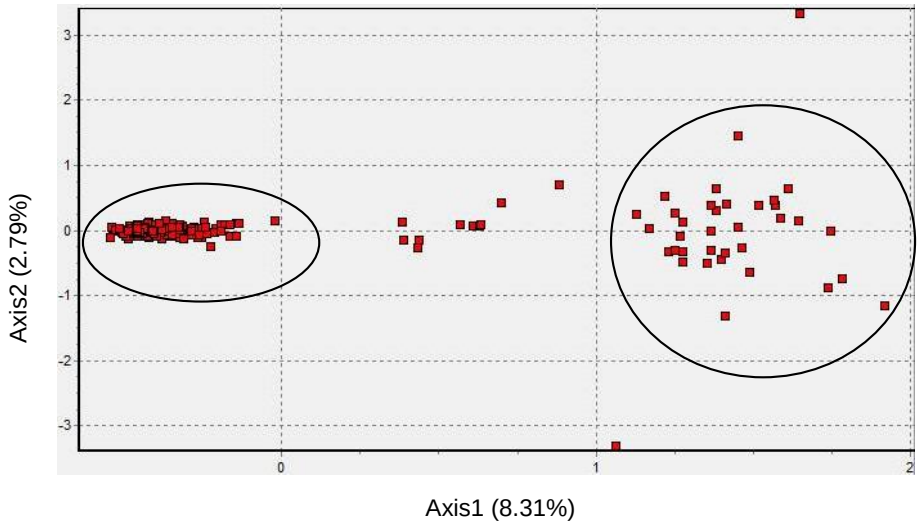


Figure 4 - Results of the Factorial Correspondence Analysis performed using GENETIX. Each square represents an individual.

In the initial runs of STRUCTURE with no prior information, two different clusters corresponding to the different species were clearly identified, indicating that the set of markers performs well for species identification. Runs performed with additional models present similar results. Also, similar results were obtained in runs where additional prior information was provided to the software by flagging populations from Madalena and Torreira as pure of each species. The vast majority of individuals from Madalena and Torreira had a high proportion of their genome (> 0.92%) assigned to *P. bocagei* and *P. carbonelli*, respectively, suggesting that in general these represent pure individuals of each species. However, two individuals from Madalena showed a low (less than 69%) but non-negligible proportion of their genome assigned to *P. carbonelli*. Runs performed with pure individuals present misleading results if used as references as the presence of admixture in two individuals could be related with inbreeding of

substructure. For this reason, we took into account the results of STRUCTURE obtained when only the individuals from the contact zone were included in the analyses.

Table 2 - Total number of individual genotypes obtained (N Total), effective number of samples assigned to each cluster in the STRUCTURE analysis (N STRUCTURE assignment) and percentage of samples assigned to each cluster in relation to the total number of samples analysed, using a threshold of 0.90.

N Total	N STRUCTURE assignment		% sampled individuals
195	*P. carbonelli*	144	73.80%
	P. bocagei	33	17%
	Hybrids	18	9.20%

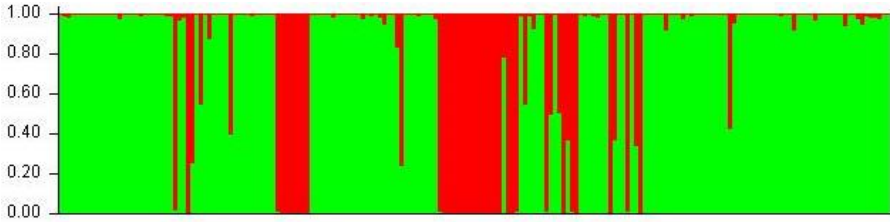


Figure 5 - Model-based multilocus genotype analyses performed in STRUCTURE showing the proportion of the genome of each genotype originating from each of the two species at K=2 (upper graph) at Espinho. Each individual is represented by a vertical bar divided in two segments, the length of which is proportional to the estimated percentage of the genome originated from *P. bocagei* (red) or *P. carbonelli* (green).

Simulated genotypes run in STRUCTURE software, reveal an average proportion of the genome assigned to the correct parental class of over 0.999 with a minimum value of 0.915 in *P. carbonelli* and an average value of 0.998 with a minimum value of 0.922 in *P. bocagei*. Therefore, a threshold of 0.90 was established for assignment procedures. Moreover, the average proportion of the genome assigned to one of the species was 0.888 (highest qi_{F1} value – 0.997) in F1 simulated genotypes and 0.713 (highest qi_{F2} value - 1) in F2 simulated genotypes. For backcrosses the average proportion of the genome assigned to each of these species were 0.811 and 0.882, respectively. It is important to notice that, despite these distinct averages, there is some overlap between the values obtained in simulated pure and hybrid individuals.

In the contact population of Espinho several individuals showed signs of admixture but the majority of the samples were attributed to one of the parental forms using the threshold of 0.90 (Figure 5). Assuming this threshold, 144 individuals were classified as *P. carbonelli*, 33 as *P. bocagei* and 18 as hybrids (Table 2), translating into a proportion of hybrids of 9.2%.

Initial runs in NEWHYBRIDS using the reference population detected a clear distinctiveness between the two species (as detected by STRUCTURE), indicating the power of our markers.

In order to help to determine a threshold for the hybrid class classification, establish the power of our markers and test the degree of confidence in hybrid assignment, NEWHYBRIDS was run using individuals of known hybrid classes generated through simulations. All pure individuals of both parental forms were detected as such by NEWHYBRIDS with more than 90% PP. Eighty-eight percent of the simulated F1 individuals were identified as such with over 75% PP (87% with higher than 90%). F2 individuals follow the same pattern with also 88% of individuals identified as such with over 75% PP (52% higher than 90%). In the backcrosses the assignment was high with 80% and 86% correctly assigned individuals for backcrosses with *P. carbonelli* and *P. bocagei*, respectively (60% and 86% with higher than 90% PP). This suggests that the used set of markers is able to correctly identify individuals belonging to both first and second generation hybrid classes. It is noteworthy that no pure individuals were classified as hybrids and vice-versa, showing that the markers used are powerful in the detection of hybridization. Also, a relatively small number of hybrids (7%) was assigned to a wrong hybrid class, lending confidence in the identification of hybrid classes in NEWHYBRIDS.

Using NEWHYBRIDS, in the contact area of Espinho we detected three F1 hybrids, four F2s, five backcrosses with *P. carbonelli* and four backcrosses with *P. bocagei*. We also detected two hybrids in which the PP of assignment were scattered throughout the four hybrid classes. These individuals are most likely not pure (according to our simulations) and could be a result of multiple generations of backcrossing (Table 3). It is noteworthy to mention that the results of NEWHYBRIDS and STRUCTURE were highly congruent in the assignment of the individuals as pure or hybrid.

With respect to mtDNA variation, in the northern region of the sampling area only mtDNA from *P. bocagei* was found while in the southern region only mtDNA from *P. carbonelli* was detected. Only in the center of the sampling area, in the 500 meters where both morphotypes are found in strict syntopy, both mtDNA types were detected, with more individuals (approximately 80%) attributed to *P. carbonelli*. The individuals of each species identified as “pure” also had the maternal lineage attributed to the same parental form, showing congruence between nuclear and mtDNA. In the case of the hybrids, all first generation hybrids and the two unspecified hybrids (that were not clearly attributed to a hybrid class) had a maternal lineage of *P. bocagei*. Among F2 hybrids, three individuals carried mtDNA of *P. bocagei* and only one of *P. carbonelli*.

The backcrossed individuals showed the maternal lineage of the respective parental form (Table 3). Globally, 12 of the 18 identified hybrids carried the maternal lineage of *P. bocagei* comprising 67% of the hybrid individuals.

Table 3 - Results obtained using NEWHYBRIDS and mtDNA depicted for all hybrid individuals. The six classes represent the proportion of the estimated posterior probabilities of assignment to pure *P. carbonelli* (PC), *P. bocagei* (PB), F1, F2, and backcross of a F1 hybrid with a pure *P. carbonelli* (BC) or with a pure *P. bocagei* (BB).

Code	mtDNA	PC	PB	F1	F2	BC	BB
19579	PB	0	0.66	0	0	0	0.33
19583	PB	0	0	0.02	0.15	0	0.84
19585	PB	0	0	0.53	0.22	0.25	0
19587	PC	0.14	0	0	0.02	0.84	0
19592	PC	0	0	0	0.94	0	0.05
19653	PC	0.88	0	0	0	0.11	0
19656	PC	0.87	0	0	0	0.13	0
19657	PB	0	0	0	0.36	0	0.63
19682	PC	0.01	0	0	0.07	0.92	0
19698	PB	0	0	0.99	0	0	0
19704	PB	0	0	0.78	0.16	0.06	0
19706	PB	0	0	0.02	0.95	0.02	0
19708	PB	0	0	0.28	0.36	0	0.36
19719	PB	0	0	0	0.88	0	0.11
19724	PB	0	0.77	0	0	0	0.23
19726	PB	0	0	0	0.83	0	0.17
19929	PC	0.51	0	0	0	0.49	0
19944	PB	0	0	0.98	0.02	0	0

2.3.3. Genetic variability in the contact zone

In *P. bocagei* from the contact zone the number of alleles ranged from 2 to 18 alleles with a mean number of 9. In *P. carbonelli* the mean number of alleles was lower (8.125) ranging from 3 to 14 alleles. *P. bocagei* presented a mean heterozygosity of 0.71 (± 0.17), higher than *P. carbonelli* with a mean of 0.63 (± 0.21) (Table 4).

In the population of Espinho (excluding hybrid individuals) loci with significant deviations from Hardy-Weinberg (HW) expectations were also found but only in *P. carbonelli* (Ph25; Ph70; Ph61). Linkage disequilibrium (LD) was also found in *P. carbonelli* in the contact zone showing association between two pairs of loci: Ph22 and Ph39, and Ph38 and Ph70. The occurrence of significant deviations from HWLE in *P.*

carbonelli at the contact zone can be a result of inbreeding or substructure in that sample, as STRUCTURE analyses for K=3 also seem to depict (see Figure S1, in Supplementary material).

All the 16 microsatellites exhibit private alleles (Pa) in at least one of the two species in the sampled area (Table 2). *P. bocagei* present more private alleles over all loci (34 Pa, distributed over 14 loci), than *P. carbonelli* (13 Pa distributed over 8 loci).

Table 4 - Number of alleles (Na), number of private alleles (Pa) and expected heterozigosity (He) observed per locus in each species in Espinho and in the total sample. Fst, Pid (probability of identity) and Pid sibling (Probability of identity in siblings) values for each locus in all analysed samples.

Locus	Na	He	*P. bocagei*			*P. carbonelli*			Fst	Pid	PidSib
			Na	Pa	He	Na	Pa	He			
Ph14	11	0.661	9	3	0.743	8	0	0.532	0.339	0.136	0.452
Ph21	25	0.878	18	5	0.901	14	2	0.830	0.121	0.021	0.314
Ph22	8	0.731	6	2	0.706	4	0	0.615	0.336	0.101	0.408
Ph25	10	0.726	8	1	0.764	8	1	0.659	0.187	0.106	0.410
Ph30	14	0.826	9	3	0.610	11	1	0.825	0.144	0.054	0.355
Ph31	16	0.887	12	2	0.885	11	1	0.859	0.081	0.021	0.311
Ph32	7	0.722	7	1	0.720	6	0	0.697	0.088	0.113	0.415
Ph38	10	0.763	7	2	0.561	6	1	0.727	0.240	0.087	0.387
Ph39	10	0.664	4	0	0.574	9	3	0.560	0.369	0.131	0.449
Ph43	5	0.379	5	2	0.325	3	0	0.111	0.827	0.383	0.638
Ph61	11	0.500	10	3	0.798	5	0	0.312	0.436	0.235	0.543
Ph62	12	0.854	9	1	0.713	11	0	0.825	0.170	0.036	0.332
Ph70	20	0.905	14	3	0.916	14	2	0.868	0.093	0.014	0.300
Ph81	16	0.712	12	1	0.880	11	2	0.578	0.287	0.092	0.411
Ph83_11	14	0.693	12	5	0.821	5	0	0.554	0.328	0.111	0.426
Ph128	4	0.537	2	0	0.451	4	2	0.548	0.042	0.325	0.566
								Overall among loci	0.257	6.77E-18	6.47E-07

The results of the AMOVA performed in Espinho (excluding hybrid individuals identified by STRUCTURE) revealed that the highest amount of variation was found within individuals (68.4%) and between populations (25.7%; Fst is significant, $p<0.001$) with the lowest amount of variation being found among individuals within populations (5.9%). The loci contributing more for the observed differentiation were Ph43, Ph61 and Ph39 with Fst values of 0.827, 0.436 and 0.368, respectively ($p<0.001$ for all these loci, see Table 4).

2.3.4. Pholidotic patterns

The distribution of character-state frequencies of the categorical pholidotic characters was evaluated in the different groups in order to investigate their variation. As all of the recorded characters were almost fixed in one of the states, no sufficient variation was available for further analyses. Ninety-seven and ninety-six percent of the individuals of Espinho showed presence of tympanic and masseteric scale, respectively. Also, in 98% of the individuals contact was present between the internasal and frontal scales while 99% do not exhibit contact between the occipital and interparietal scales. One of the supposedly continuous pholidotic characters, the number of supralabial scales presented less than five categories and was therefore excluded from ANOVA analyses since it did not conform to a continuous distribution. Therefore, these five traits were not used in subsequent morphological analyses.

Analysis of variance (ANOVA) for each of the continuous pholidotic characters recorded indicated a significant effect of species and sex (Table 6). Interaction terms were not significant in any of the studied characters. In most characters, males from both species showed higher scale values than females. The only exception was VSN, which was higher in females of both species. Comparing the two species, both sexes of *P. carbonelli* present more femoral pores (FPN), supraciliary granules (SCGN) and supratemporal scales (STSN) and less supralabial scales (SLSN) than *P. bocagei* (Table 5).

Table 5 – Mean value and standard error of each variable analysed in *P. carbonelli* and *P. bocagei*. See methods for variable abbreviations.

Variable	***P. carbonelli***		***P. bocagei***	
	Males	Females	Males	Females
FPN	18.03±0.18	17.14±0.19	16.71±0.37	16.80±0.40
CSN	9.69±0.12	9.54±0.13	9.64±0.26	9.67±0.28
GSN	26.23±0.33	26.25±0.36	26.36±0.70	26.33±0.75
VSN	27.02±0.18	29.46±0.20	27.21±0.38	30.08±0.41
SCSN	5.41±0.10	5.38±0.11	5.64±0.20	5.17±0.22
SCGN	9.40±0.26	9.69±0.29	7.61±0.55	7.58±0.59
STSN	3.80±0.11	3.71±0.12	3.43±0.23	3.08±0.25
SLSN	6.43±0.07	6.21±0.08	6.64±0.15	6.50±0.16
SDLN	24.84±0.21	23.79±0.23	24.93±0.43	23.75±0.47
N	61	52	14	12

Table 6 - Results of ANOVA showing the p- values at the factors sex and species and the respective interaction term for each morphological trait used. See methods for variable abbreviations.

	FPN	CSN	GSN	VSN	SCSN	SCGN	STSN	SDLN
Sex	0.013	0.767	0.998	0.000	0.133	0.767	0.243	0.002
Species	0.000	0.846	0.854	0.192	0.964	0.000	0.008	0.939
Sex*Species	0.671	0.683	0.969	0.499	0.176	0.726	0.498	0.854

Three characters showed significant differences between sexes (FPN, VSN, and SDLN, Table 6). Significant differences between species were only detected in the traits FPN, SCGN and STSN. No significant differences were detected at the interaction term. The results of Principal Component Analysis (PCA) indicate that global pholidotic variation between sexes mainly results from VSN, FPN and SDLN. The factors that contribute the most to the variation between species are SCGN, FPN and STSN (Table 7), matching once again the significant differences detected by the ANOVA.

Table 7 - Correlations between the first four principal component axes and initial variables as obtained from the principal component analyses (PCA) applied to continuous pholidotic variables in sex and species. %exp.: percentage of variation explained by the each axis; %Cum.: cumulative percentage explained. See methods for variable abbreviations

	Sex				Species			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
FPNM	0.488	0.354	0.237	-0.188	-0.456	0.445	-0.319	0.030
CSN	0.129	0.177	0.476	0.754	0.008	0.378	0.100	0.760
GSN	0.145	0.640	0.456	-0.138	-0.215	0.705	0.221	0.049
VSN	-0.737	0.288	0.230	0.026	0.059	-0.173	0.619	0.348
SCSN	0.147	0.528	-0.384	0.185	-0.178	0.320	0.484	-0.398
SCGN	0.019	0.588	-0.189	-0.090	-0.539	0.114	0.302	-0.010
STSN	0.164	0.552	-0.207	-0.321	-0.449	0.284	0.217	-0.397
SLSN	0.248	-0.354	0.485	-0.499	0.294	0.198	-0.593	-0.112
SDLN	0.485	0.164	0.523	-0.041	-0.021	0.693	-0.317	0.111
% exp.	26.94	16.04	12.06	9.45	22.62	16.08	13.20	9.74
% Cum.	26.94	42.97	55.03	64.49	22.62	38.70	51.90	61.64

Finally, in the PCA performed using only the three characters that exhibited significant differences between species (FPN, STSN and SCGN) we observed extensive overlap among individuals of the two parental forms and the presence of hybrids in all the range of the axes (Figure 6).

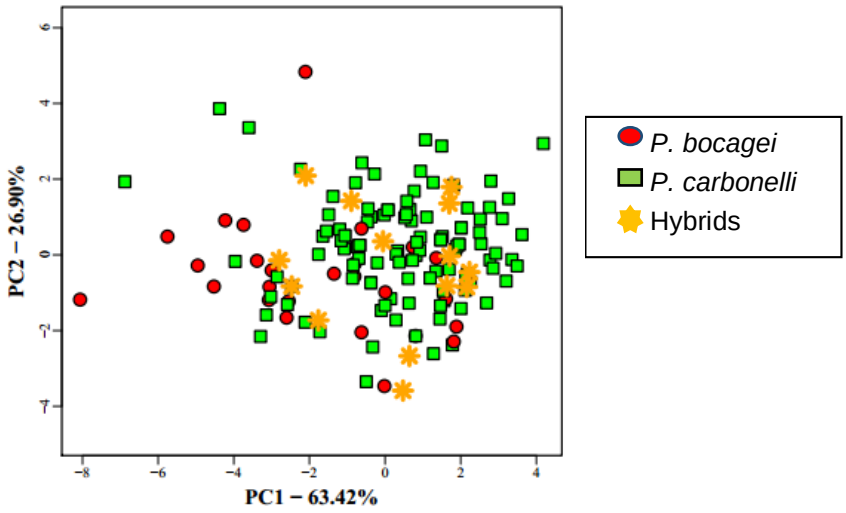


Figure 6 - Scatter-plots of individual scores of the first two principal components. The x axis represent 63.42% of the variance and the y axis 26.90 %. Analysis performed with the three pholidotic characters that present significant differences between species (FPN, STSN and SCGN).

2.3.5. Discriminant analyses

The main objective of the analysis was to obtain a reliable PP of the hybrid individuals belonging to each of the parental species on the basis of morphology.

The ANOVA comparing allopatric localities with the pure individuals from Espinho showed significant differences in both species. However, when analysing each character separately only the supratemporal scales showed significant differences between sites in both species. As we only observed significant differences in only one of the characters used, we proceeded in using all of them for constructing the discriminant function to be implemented for examining the morphological properties of the hybrid zone. Therefore, we were able to use allopatric populations in the construction of the discriminant function.

The percentage of correct classification using only reference samples (allopatric populations) was 80% in *P. bocagei* and 78% in *P. carbonelli*. The level of correct classification was lower when the total cross validation sample (both allopatric and hybrid zone individuals) was taken into account, as the percent of correct classification was 77% in *P. bocagei* and approximately 74% in *P. carbonelli*. Within the cross-validation sample, individuals from the hybrid zone show lower levels of correct classification (63% in pure *P. carbonelli* and 31% in pure *P. bocagei*) indicating a higher morphological proximity of the two species in Espinho.

The application of the discriminant function to hybrid individuals allowed an evaluation of their morphological assignment to each of the species. Table 8 shows the

posterior probabilities obtained with the application of the discriminant function based on pholidotic traits for each of the hybrid individuals identified by genetic analysis using the discriminant function constructed using both allopatric and hybrid zone individuals. Similar results were obtained when using the discriminant function constructed using allopatric populations only. Hybrids seem to be morphologically attributed to both parental species in a similar ratio and no clear relation was detected between the hybrid classes determined from genetic markers and morphological assignments.

Table 8 - Posterior probabilities of belonging to *P. bocagei* (PB) or *P. carbonelli* (PC) based on the discriminant function performed with morphological data presented for all hybrid individuals detected with genetic markers and classified according to the results of NEWHYBRIDS: class classification (F1, F2, and F1 backcrossed with *P. bocagei* (BB) or *P. carbonelli* (BC)).

Code	**NewHybrids**	**Sex**	**PB**	**PC**
19698	F1	F	0.68	0.32
19704	F1	F	0.83	0.17
19944	F1	F	0.11	0.89
19592	F2	F	0.65	0.35
19706	F2	F	0.64	0.36
19719	F2	M	0.56	0.44
19726	F2	F	0.21	0.79
19587	BC	M	0.21	0.79
19653	BC	F	0.42	0.58
19656	BC	M	0.85	0.15
19682	BC	F	0.24	0.76
19929	BC	F	0.82	0.18
19579	BB	F	0.14	0.86
19583	BB	F	0.36	0.64
19657	BB	M	0.31	0.69
19724	BB	F	0.42	0.58
19585	Unsp.	M	0.14	0.86
19708	Unsp.	M	0.76	0.24

This result is confirmed when the proportion of the genome originating in each species obtained estimated using STRUCTURE based on the genetic data and the PP of assignment obtained using morphological data were compared (Figure 7). Indeed, the two species are well separated when considering their genetic assignment (x axis), with the hybrids located in the middle, showing the power of our genetic markers. By contrast, a totally different scenario is observed when examining the morphological assignment (y axis), as both parental species and hybrids are located across the entire range. A significant correlation between the morphological and genetic assignment was found for samples of *P. bocagei* (Spearman's R = 0.366, p = 0.02) but not for samples of *P. carbonelli* (Spearman's R = 0.079, p = 0.35).

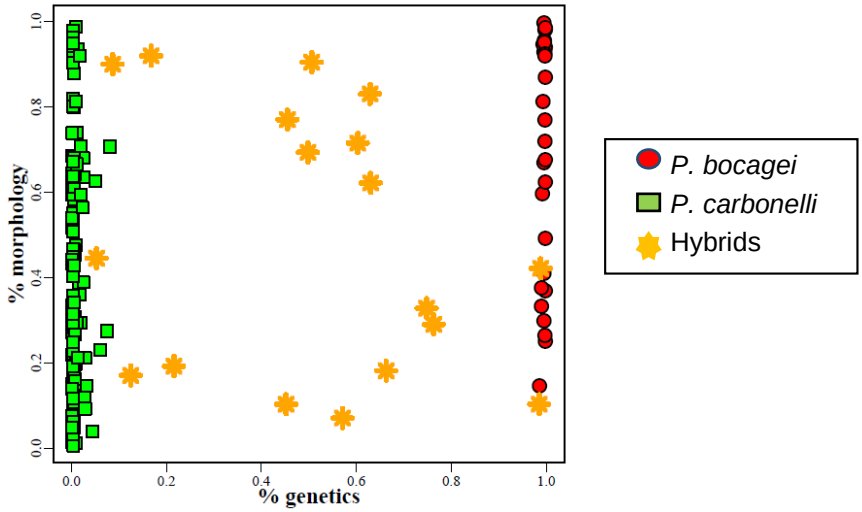


Figure 7 - Scatter-plot of the genetic versus the morphological assignment to *P. bocagei* for the individuals sampled in the hybrid zone.

2.3.6. Asymmetry analyses

The two-way ANOVAs show significant effects ($p<0.05$) of the factor Individual and of the interaction term (Individual*Side) in all the groups for FPN and SCGN. This indicates the absence of directional asymmetry in these traits and the presence of fluctuating asymmetry in both pure and hybrid individuals. Measurement error was significantly lower than between side variation. In the case of SDLN the same pattern was detected except in *P. bocagei* where the results do not support fluctuating asymmetry but rather directional asymmetry ($p<0.05$ in Side factor, Table 9).

Table 9 - Results obtained from two-way, mixed model ANOVAs (side = fixed factor, individual = random factor) for all groups and traits, separately. df - degrees of freedom; F - F-statistic; P - corresponding p-value. See methods for variable abbreviations.

		Individual effect			Side effect			Individual*Side		
Group	**Trait**	df	F	P	df	F	P	df	F	P
Hybrids	FPN	17	11.25	<0.001	1	1.35	0.26	17	59.82	<0.001
	SCGN	17	4.29	<0.001	1	0.57	0.46	17	4.58	<0.001
	SDLN	13	1.72	0.17	1	0.69	0.42	13	3.46	<0.001
P. bocagei	FPN	31	10.52	<0.001	1	0.22	0.65	31	115.83	<0.001
	SCGN	32	10.6	<0.001	1	0.64	0.43	32	1.94	0.01
	SDLN	22	11.48	<0.001	1	7.99	0.00	22	1.36	0.18
P. carbonelli	FPN	139	10.54	<0.001	1	3.05	0.08	139	35.65	<0.001
	SCGN	139	4.59	<0.001	1	1.85	0.18	139	3.23	<0.001
	SDLN	103	7.79	<0.001	1	0.37	0.54	103	3.85	<0.001

None of the two AI (AI1 and AI2) showed significant deviations from normality for any of the groups/traits using a Kolmogorov-Smirnov test. This allowed discarding the existence of antisymmetry in our sample. The test of total body size and trait dependence through linear regression of |R – L| on SVL and (R+L)/2 showed no size dependence for any of the traits/groups (Table 10).

Table 10 - Results obtained from linear regression of |R-L| on SVL and (R+L)/2 for all traits. Df - degrees of freedom; F - F-statistic; P - corresponding p-value. See methods for variable abbreviations.

Trait	**SVL**			**(R+L)/2**		
	df	F	P	df	F	P
FPN|R-L|	1	0.511	0.475	1	1.003	0.317
SCGN|R-L|	1	0.392	0.531	1	0.168	0.681
SDLN|R-L|	1	0.138	0.709	1	0.549	0.459

As all the groups exhibited fluctuating asymmetry at two of the traits, the ANOVA design taking into account the factor type and sex was used in order to compare levels of FA in the parental forms (*P. bocagei* and *P. carbonelli*) and hybrid individuals. The results of this analysis do not support the existence of significant differences between the hybrids and parental forms (Factor type), thus not supporting increased asymmetry in hybrid individuals. Also, no significant differences in FA were found between sexes (Table 11).

Table 11 - Results obtained with three-way ANOVA on log transformed average of trait values across the two replicate counts, with sex, type (Parental vs. Hybrids), species nested within type and trait as factors and all interaction effects.

		SS	**df**	**F**	**P**
FPN	**Type**	0.000012	1	0.006625	0.93522
	Species(Type)	0.002755	2	0.737659	0.479672
	Sex	0.000277	1	0.148127	0.700786
	Error	0.336103	180		
SCGN	**Type**	0.491868	1	0.008995	0.924544
	Species(Type)	11.72065	2	0.107169	0.89843
	Sex	44.01634	1	0.804937	0.370794
	Error	10061.66	184		
SDLN	**Type**	0.000066	1	0.045953	0.830594
	Species(Type)	0.006453	2	2.233472	0.111221
	Sex	0.000678	1	0.469237	0.494549
	Error	0.189248	131		

2.4. Discussion

2.4.1. Genetic characterization of the hybrid zone

The number of pure individuals in our sample is much higher than the number of hybrids, a result that supports previous studies (Pinho *et al.* 2009) that reported this area to be bimodal hybrid zone (Jiggins & Mallet 2000). The fact that the majority of the sampled individuals do not show signs of genetic admixture despite the broad co-existence of the two species supports the existence of strong barriers to gene flow as previously suggested. However, an important difference from the previous analysis was the finding of F1 individuals and other hybrids presenting similar proportions of the genome originating in both species. This hybrid zone may therefore also approximate the definition of a trimodal hybrid zone as suggested by Gay *et al.* (2008), although in this case the difference in frequency between the modes are extreme.

In 1922 J. B. S. Haldane (Haldane 1922) postulated that in first generation hybrids, when one sex in a population of hybridizing taxa is absent, rare or sterile, it is generally the heterogametic sex. This rule has been demonstrated in a series of different animal groups, suggesting that speciation by postzygotic isolation occurs in a similar way among very different groups of animals such as mammals, birds, insects (Orr 1997) and lacertid lizards (Murphy *et al.* 2000). In lizards of the genus *Podarcis*, as in the majority of the lacertids, females are the heterogametic sex (Olmo 2005). Therefore, if this rule applies to the hybrid zone between *P. bocagei* and *P. carbonelli* we would expect F1 females to be absent due to female unviability. Furthermore, we would expect backcrosses to one of the parental species to present only mtDNA of that species as the mating generating those backcrosses would only be able to occur between hybrid males and pure females. In our sample the three F1 individuals detected were females, thus allowing us to reject complete female unviability. However, all backcrosses with *P. carbonelli* (5 individuals) and with *P. bocagei* (4 individuals) detected in this study present mtDNA strictly from that species. Despite the low sample size, these results suggest the possibility of reduced F1 female fertility. Previous studies performed in this hybrid zone found no F1 hybrids, thus rendering impossible to test for female unviability. Nevertheless, they reported one backcross to *P. bocagei* carrying *P. carbonelli* mtDNA (Pinho *et al.* 2009). Together, these results suggest a reduced, but not total, viability or sterility of F1 females in this hybrid zone. It would be interesting to compare survival and reproductive success of F1 females with those of F1 males, by performing controlled tests in the laboratory, as it was already show that hybrids could be originated in captivity (Galán, 2002; Pinho *et al.*, unpublished data).

Approximately 67% of the hybrids, including the three F1 individuals, had a mtDNA haplotype originated from *P. bocagei*, clearly contrasting with the much lower frequency of this mitotype in the overall sample. Although more data is clearly necessary to ascertain the significance of this trend, these results suggest an asymmetry in the direction of hybridization towards females of *P. bocagei* copulating with males of *P. carbonelli* and hybrid males. As *P. bocagei* is larger in body size than *P. carbonelli*, this preference of males towards females of *P. bocagei* may have been favoured by their higher expected reproductive success: bigger females produce more clutches and lay more eggs per clutch than smaller females (Carretero & Llorente 1993; Galán 1999; Fitze *et al.* 2005). Larger females also have a tendency to be more promiscuous, ensuring better offspring (Olsson *et al.* 1994). Another possible explanation would be the existence of some type of asymmetrical incompatibilities between the two species. Although not testable in the framework of this study, this bias provides an interesting working hypothesis for future studies.

2.4.2. Morphological characterization of the hybrid zone

Morphological studies are extremely important for understanding the evolutionary mechanisms involved in phenotypic evolution during species differentiation (Adams *et al.* 2009). In the present study we studied and compared the morphological characteristics of parental forms and hybrids by quantifying characters normally used in taxonomic studies in the Lacertidae.

Results of the analyses of pholidotic patterns it clearly show that sexual dimorphism is present in both species as three characters display significant differences between sexes. Males of both species present more femoral pores and subdigital lamellae than females and females present more ventral scales. Our results are similar to the those presented by Kaliontzopoulou *et al*. (2005). Normally, femoral pores show sexual dimorphism in most lizards (Cole 1966) with males usually displaying a higher number of femoral pores (Carretero & Llorente 1993; Carretero *et al.* 2003). In the ventral scales the higher trait value in females must be related with longer trunks which are needed to the allocation of eggs (Carretero & Llorente 1993; Braña 1996). As a more detailed study of sexual dimorphism is out of the scope of this study no further details will be discussed.

Only three of the pholidotic characters analyzed (FPN, STSN and SCGN) presented significant differences between species. The PCA using these characters (Figure 6) suggests a large overlap between species, a usual pattern in Iberian *Podarcis* species (Kaliontzopoulou *et al.* 2005, 2012b). This difficulty in the

classification of the individuals is probably related with the high intraspecific variability, typical of the *Podarcis* genus, which tends to difficult statistical diagnosis (Kaliontzopoulou *et al.* 2012b). Using only characters with significant differences between species, we were able to obtain a discriminant function that presented a relatively high percentage of correct classification either using only reference populations (around 80% in both species) or using allopatric and hybrid-zone individuals (77% in *P. bocagei* and 74% in *P. carbonelli*). These results confirm the morphological differentiation between these species. We should also stress that the species morphological assignment done in the field showed higher rate of correct classification (89% of correct classification in *P. bocagei* and 96% in *P. carbonelli* identified according to genetic data) than the one obtained with the DA (results not shown). These results suggest that reliable morphological discrimination between species can be accomplished when the right framework and traits are used.

Contrasting to the high classification scores of individuals from the reference populations, individuals from the hybrid zone present a very low percentage of correct classification (63% in *P. carbonelli* and 31% in *P. bocagei*), indicating higher similarity between the species in this area. This similarity may indicate that pholidotic patterns are not only a result of evolutionary trajectories but are also influenced by environmental and ecological factors. Several hypotheses can be put forward to explain the apparent morphological convergence between the two species observed in the hybrid zone: *i)* phenotypic plasticity*, ii*) differences in selective regimes across the species' ranges and/or similar patterns of local adaptation in the contact zone in both species (Martínez-Freiría et al., 2008 and references therein), *iii)* admixture and introgression between the two species. However, in order to test these hypotheses additional experiments are required (such as common garden experiments with both species). In any case, this result suggests that morphological distinctiveness is not increased in the contact zone, as would be expected if a process of reinforcement of reproductive isolation was taking place.

The major aim of the study of pholidotic patterns was compare morphological and genetic assignments of every single individual to investigate the influence of hybridization on the phenotype. A first observation is that hybrids are clearly not intrinsically different, morphologically speaking, from each of the parental forms, i.e. there is no such thing as a typical hybrid morphology. Moreover, according to the posterior probability of assignment, hybrids were equally probable to resemble *P. bocagei* or *P. carbonelli* and are not more similar to one of the species (Table 8). At the same time, no relation was found between genetic and morphological assignments when the posterior probabilities and proportion of the genome originating in each

species were compared (Figure 7). It is important to stress that this lack of correlation between genetic and morphological assignments can be explained by the low number of pholidotic characters used.

Another aspect of hybrid morphology that we assessed was fluctuating asymmetry. The results of this analysis do not support the hypothesis of increased levels of fluctuation asymmetry in hybrids, as these individuals do not display higher asymmetry than individuals from the hybrid zone genetically assigned to parental forms. Although previous studies (Palmer & Strobeck 1986; Palmer 1994; Lazić *et al.* 2013) suggest the importance of using multiple traits and populations when testing for fluctuating asymmetry, hybridisation between these two species is only known to occur in a single locality. Also, the low number of hybrids detected in our sample could reduce the power of statistical analyses to detect significant differences between groups. Accordingly these results should be taken with caution.

2.4.3. Evaluating the stability of hybrid zone inferences

When our results are compared to a previous study of this hybrid zone, which used samples collected between 2001 and 2002 (Pinho et al., 2009), important congruences and differences are found, that may allow us to better understand the evolutionary and demographic dynamics of this hybrid zone.

As in the previous study, we detected ongoing hybridization and roughly around the same area as before, which indicates some stability in hybrid zone dynamics. Our samples also display a highly bimodal distribution of genotypes, supporting the existence of strong barriers to gene flow preventing free admixture between these two species. However, there are important differences that may not be easily explained: i) we detected a much lower proportion of hybrids (9.2%), than that documented in the previous study (~30% according to an inspection of Figure 3 in Pinho *et al.* 2009); ii) we detected a higher percentage of hybrids with intermediate proportion of the genomes originating in the two species, including three F1s; and finally iii) we report a bias towards the *P. bocagei* mitotype in hybrids which had not been previously detected.

One possible explanation for the discrepancies in the proportion of hybrids lies in differences between the two studies in terms of the size of sampled area in the contact zone (much larger in the present study). In fact if the area of hybridization is restricted to a small portion of the sampled area (as it seems to be, see Chapter 3), extending the sampled area would result in an increased sampling of pure individuals. However, considering only the area of strict morphological overlap where the previous

samples were collected, the percentage of hybrids is still much lower. Moreover, this fact alone does not account for the other differences in admixture patterns.

Another possible explanation for these discrepancies is the different sampling strategy employed in both studies. Whereas Pinho *et al.* (2009) used an artificially balanced sample between the two species, forcing the sample to be roughly composed by 50% of each species, we sampled, as much as possible, in proportion to the actual frequency of the species in the contact area since we collected as many lizards as we could.

Finally, a possibility that cannot be disregarded is the change in introgression patterns due to ecological, demographic or evolutionary changes over the years. In fact, 11-12 years (possibly corresponding to an equal number of lizard generations), separate the two sampling events. It is possible that slight changes in the location of the contact zone (with corresponding shifts in the relative frequencies of the two species and hybrids) may occur within this time frame simply by stochasticity. During this period, the patterns of selection acting on both species have also probably changed. For example, ecological changes in this are a possibly due to anthropogenic activities and climatic phenomena may help explain some of the intriguing differences reported in this study.

Most of these hypotheses imply that taking into account possible spatial segregation may help reconcile different patterns observed by the different studies. Therefore, our results also highlight the importance of performing additional work to further characterise and understand this interesting hybrid zone. In this context, spatial analyses would be particularly interesting as these would provide information on the geographic distribution of the specimens and alleles at this micro-scale allowing to infer the role of intrinsic and extrinsic factors in the restriction of gene flow between these two species.

3. Zooming in a hybrid zone: the micro-spatial structure of the contact zone between two species of wall lizards

3.1. Introduction

Speciation can be defined as the development of mechanisms responsible for reproductive isolation. One of the best methods to study these barriers is to investigate cases in which they are not completely effective. In hybrid zones the effect of natural selection in maintaining the integrity of the parental forms is destabilized by the presence of intermediate individuals and the consequent homogenizing effects of migration and recombination (Barton & Hewitt 1985). At the same time, the maintenance of the parental forms without complete reproductive isolation provides an interesting scenario to study the mechanisms preventing gene flow and to gain insights to the mechanisms driving speciation. It is for this reason that many questions about speciation have been addressed in studies performed in hybrid zones (Barton & Hewitt 1985; Jiggins & Mallet 2000; Abbott *et al.* 2013).

The balance between hybridization and the lower fitness of hybrids when compared to parental forms is what maintains a hybrid zone stable (Tarroso *et al.* 2014). Reduced hybrid fitness can be due to endogenous selection, which is independent of environmental factors (Barton & Hewitt 1989), or due to exogenous selection, which is caused by the hybrids' maladaptation to the environment (Barton & Hewitt 1985; Doebeli & Dieckmann 2003; Gifford 2008; Tarroso *et al.* 2014). Studies on hybrid zones provide examples of the co-occurrence of both types of selection promoting species divergence (Ross & Harrison 2002; Phillips *et al.* 2004; Tarroso *et al.* 2014).

From a geographical perspective, hybrid zones could be structured in two different ways. First, genotypes and phenotypes may vary following a more or less smooth cline with a transition from one parental form to another as narrow areas of overlap or narrow contact zones (Barton & Hewitt 1985). Alternatively, they may present a mosaic structure with a patchy/mosaic distribution of genotypes and phenotypes across the landscape, a spatial pattern which is usually related with environmental heterogeneity (Britch *et al.* 2001). The study of the spatial distribution of

genotypes and phenotypes across a hybrid zone has always been an important subject for evolutionary ecologists that want to investigate the processes generating and maintaining hybrid zones (Barton & Hewitt 1985; Benson *et al.* 2012), as understanding the causes of the spatial patterns detected help to identify the nature and strength of reproductive barriers promoting isolation. Therefore it is important to combine population genetic analysis with the location of genotypes in a spatially explicit framework (Schoville & Bonin 2012). One possible way of doing this is by modeling the shape of clines of genetic and/or morphological traits. Clines represent changes in character frequencies along a geographic transect. Divergence between populations/species result in character frequency differences between them. When species hybridize after the secondary contact, the shape of the clines, as well as the concordance of clines from different loci, varies according to the strength of selection against hybrids. In a "tension zone" (Barton & Hewitt 1985), the position and width of clines are maintained by a balance between dispersal into and from the contact area which has a homogenizing effect and selection against hybrids which opposes to homogenization (Barton & Hewitt 1985; Kruuk *et al.* 1999; Gay *et al.* 2008). Due to hybridization, recombination breaks down allelic associations, allowing distinct genomic regions to be differently exchanged between hybridizing taxa. The center of the hybrid zone will act as an effective barrier to the dispersal of neutral and negatively selected alleles when selection against hybrids is strong, but advantageous alleles could still cross that barrier, although with some delay (Piálek & Barton 1997). Therefore, some regions do not migrate from one taxon to the other, while other regions introgress more freely (Barton 1979; Rieseberg *et al.* 1999). A comparison of the patterns of introgression among loci will therefore provide useful information to identify genomic regions associated with reproductive barriers and to evaluate the strength of selection against hybrids: if selection is weak, different traits may present different clinal patterns; if selection is strong, epistatic interactions between loci will make the genome behave as a single cohesive unit (Kruuk *et al.* 1999).

Another perspective that may be taken into account when performing hybrid zone analyses is that of landscape genetics, which also combines population genetics and spatial data to study the interaction between population dynamics and ecological factors in a spatially explicit framework (Manel *et al.* 2003). This allows addressing the interaction between landscape and microevolutionary processes such as gene flow, genetic drift and selection (Manel *et al.* 2003). One particularity of this approach is the study of evolutionary processes at the individual level. When applied to hybrid zone studies, this perspective thus has the advantage of allowing an evaluation of the distribution of individual genotypes in space, providing information about the clinal or

patchy nature of the hybrid zone and relating such distribution with particular habitat features (e.g. Shurtliff *et al.* 2014; Tarroso *et al.* 2014).

One of the important characteristic of spatially explicit frameworks such as clinal analyses and landscape genetics is scale. The spatial scale is defined by the species-specific biological and ecological processes under study, and by the spatial dimension at which we are able to sample (Holderegger & Wagner 2006).

In previous analyses (Chapter 2) we studied the genetic and morphological composition of the hybrid zone between two species of wall lizards (*P. bocagei* and *P. carbonelli*). The results obtained indicate the existence of a bimodal hybrid zone. However, the geographic distribution of genotypes and phenotypes at this geographic scale is still unknown. To overcome this lack of information and gain further insights to the possible mechanisms maintaining the identity of both species in their contact area, we perform here a fine-scale spatial analysis of the hybrid zone at the individual level in order to characterize the distribution of genotypes and phenotypes across the landscape and seek for possible factors affecting such distribution. This analysis may contribute to understand the influence of extrinsic and/or intrinsic factors shaping the hybrid zone. Specifically, we want to: 1) assess if morphological and genetic traits are spatially structured in the area where *P. bocagei* and *P. carbonelli* hybridize; 2) test if traits change according to a clinal model or are better explained by a random or a mosaic distributions; and 3) test if distinct traits present coincident patterns of frequency change across the landscape.

3.2. Methods

In order to perform a fine-scale spatial analysis of the hybrid zone between *P. bocagei* and *P. carbonelli* we recorded the precise coordinates (WGS1984 datum) using a GPS with an error of approximately 4 meters of each individual sampled during fieldwork. In the present study we use the same individuals used in previous analyses (Chapter 2), for which the proportion of the genome assigned to each of the parental species was inferred using STRUCTURE based on 16 microsatellite loci. We also assessed the mitotype carried by each individual (see Chapter 2). Also, a discriminant analysis based on morphological traits was used to calculate the posterior probabilities of each individual to belong to each of the parental forms (see Chapter 2).

To obtain a first overview of the distribution of individuals across the hybrid zone, we produced a map using only field identifications based on various empirical

diagnostic characters which include size, colour pattern and head shape (Pinho *et al.* 2009). Then, to examine the spatial localization of hybrids and compare it with the position of both parental forms we generated a second map based on the proportion of genome assigned to a given species by STRUCTURE, depicting individuals classified as pure *P. bocagei*, or *P. carbonelli* (those having at least 90% of their genome assigned to one of these species) and hybrids (those with less than 90% of their genome assigned to a species, see Chapter 2 for details on hybrid identification). Both maps were produced using QUANTUM GIS (QGIS) 2.4. (QGIS Development Team 2014).

3.2.1. Interpolations

In order to investigate the spatial distribution of genetic and morphological traits in the contact zone between the two wall lizard species, we used interpolations. Specifically, we used the inverse distance-weighted ($w=1/distance^2$) algorithm (IDW) to produce a continuous surface of variation in the traits of interest over the study area. This method estimates the value of an attribute at any given point using a linear combination of values at sampled points, weighted by an inverse function of the squared distance from the point of interest to the sampled point. The general premise of this method is that sampled points closer to the unsampled points are more similar in their values than those further away, i.e. that trait similarity varies as a squared function of geographical distance (Li & Australia 2008; Lu & Wong 2008). All interpolations were performed using QGIS 2.4. (QGIS Development Team 2014).

Genetic data

To investigate the spatial distribution of genetic composition across the hybrid zone we used two different types of data for interpolations. First, an interpolation was performed using the proportion of genome assigned to *P. carbonelli* or *P. bocagei* (given by STRUCTURE) obtained for each individual (Chapter 2). Second, the pairwise genetic distances between the individuals of the sampled area were calculated trough Nei´s genetic distance using GENALEX 6.5 (Peakall & Smouse 2006, 2012) using the genotypes of the individuals for 16 microsatellite loci (Chapter 2). The distance matrix was decomposed into linear components using principal coordinates analysis (PCoA), in order to obtain a summary linear representation of the dissimilarities between individuals. The first axis of the PCoA explains 34.21% of the variance in our sample and the second axis 7.48%. Therefore, the first axis of the PCoA was used in the

interpolation in order to visualize how these dissimilarities were distributed across the landscape of the hybrid zone.

Morphological data

To visualize patterns of geographic variation in continuous pholidotic characters, we also used two different methods, to match those used for genetic data and to be able to compare spatial patterns between genetic and morphological characters in a common spatial framework.

First, for each individual, the posterior probability (PP) of assignment to each of the parental forms (*P. carbonelli* and *P. bocagei*) was calculated using the discriminant function constructed with "pure" individuals based on morphological traits (see Chapter 2). These posterior probabilities were used to perform an interpolation that could be compared with the interpolation map produced with proportion of genome assigned to *P. carbonelli* or *P. bocagei* (see 3.1.2).

Next, in order to establish the degree of morphological similarity between individuals, we calculated multivariate Euclidian distances (ED) based on the three pholidotic characters that exhibited significant differences between species (FPN, SCGN, STSN – see Chapter 2). The resulting matrix was decomposed using PCoA and the first axis (that explains 63.42% of the variance in the sample) was used for interpolation. This interpolation could be compared with that conducted using genetic distances between individuals of Espinho.

3.2.2. Cline analyses

Although cline analyses are usually performed using localities as sampling units, an individual-based sampling is also amenable to this type of assessment. Although our samples were collected along a narrow stretch of coastline, they did not form a strict unidimensional transect, as required for cline analyses performed with the software used. Hence, in order to fit our sampling points to this type of transect, we used the projection of each individual's geographical coordinates onto the linear regression line obtained from all sampling points. The distance from the first (northernmost) individual point along this line was then taken as each sample's position in the transect.

We conducted cline analyses using two types of data: i) STRUCTURE proportions of assignment based on microsatellite data; ii) allele frequencies at each locus,

including mitochondrial DNA and microsatellites. The latter analyses require an a priori assignment of each allele as characteristic of one of the species. This is straightforward with respect to the mtDNA since both species exhibit clearly distinct lineages (e.g. Kaliontzopoulou *et al.* 2011), but the microsatellite loci used in this work did not showed diagnostic alleles between species. Hence we made this assignment based on the species in which alleles appeared at higher frequencies, based on the estimated allele frequencies in each cluster given by STRUCTURE. All *P. bocagei* and all *P. carbonelli* alleles were pooled, simulating a two-allele system for all loci: allele frequencies of each individual were 1 when both alleles were allocated to *P. carbonelli*; 0.5 when one allele was attributed to one species and another to the other one; and 0 when both alleles were attributed to *P. bocagei*. Obviously, for mtDNA only 0 and 1 frequencies were possible.

Allele frequencies for each locus and STRUCTURE proportions of assignment were then fit to a series of equilibrium geographic cline models (Szymura & Barton 1986, 1991; Gay *et al.* 2008) using the Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm implemented in the R package HZAR (Derryberry *et al.* 2014). Fifteen separate models, representing three possible combinations of trait interval (the maximum [pMax] and minimum [pMin] gene frequency values at the tail ends of a cline: i) fixed to 0 and 1; ii) observed values; iii) estimated values) and five possible combinations of fitting tails (exponential decay curves: i) none; fitted; ii) left only; iii) right only; iv) mirror tails; v) both tails estimated separately) were tested. For all models cline center (distance from the northern sampling location) and width (1/maximum slope) were estimated. We checked for convergence by repeating the run using different random seeds and checking for appropriate parameter trace and density plots. The fit to the data of 15 cline models, plus a null model representing the absence of a cline in allele frequencies were then compared using Akaike's information criterion corrected for small sample size (AICc). The confidence interval between the parameters center and width was used in the coincidence of the clines (looking for a common center) and for a common width in order to see the concordance of the 16 microsatellites, the mtDNA marker, and the STRUCTURE proportion of assignments (Chapter 2).

The R package Tcl/Tk (Grosjean 2014) was used to plot multiple clines in the same image in order to visualize the similarities or differences between clines.

3.3. Results

3.3.1. Distribution of parental and hybrid individuals across the hybrid zone

The dune area of Espinho sampled has approximately 2.3 kilometers long from north to south. The individuals are restricted to a narrow strip located in the dune, due to the presence of a railroad and no individuals were found nearby the dune area. Notice that the resulting distribution of sampled individuals is therefore not due to the sampling being performed in a North-South transect as it may be suggested by the map (Figure 8). Instead, it represents the actual distribution of individuals in the dune area of Espinho and its surroundings, as we performed a random sampling across the whole area capturing every individual sighted.

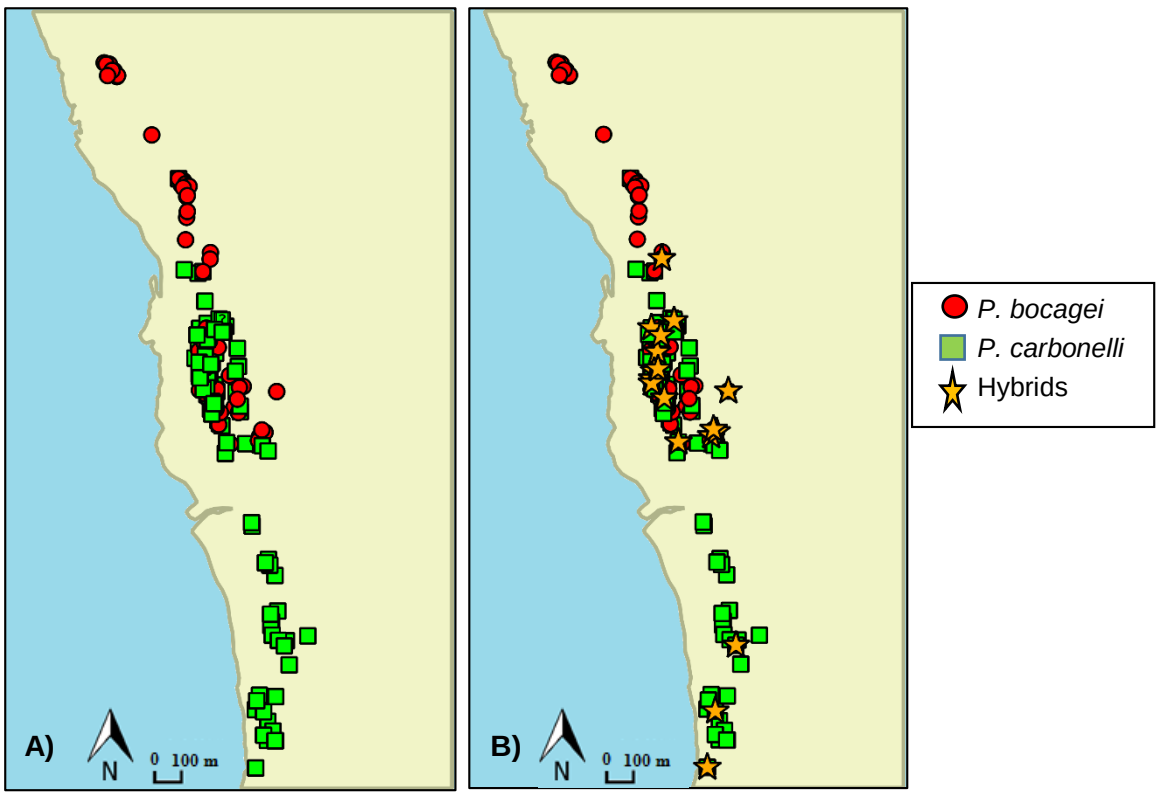


Figure 8 – Map of the study area representing the exact position where each individual was sampled. A) Individuals identities based on field identifications; B) Individuals classified based on the proportion of genome assigned to a species estimated using STRUCTURE.

Across the contact zone of Espinho, *P. bocagei* is clearly more limited to the Northern part, whereas *P. carbonelli* is mainly located in the southern part (Figure 8A). We also identified the existence of a reduced area of syntopy at the microscale, where *P. carbonelli* seems to be more frequent. Notice that it was not possible to identify

hybrid individuals by visual inspection in the field, which suggests that hybrids are similar to either one or the other parental species. Also, in a discriminant analysis (Chapter 2) using morphological traits hybrid individuals were found to be similar to one of parental forms.

Based on the genetic assignment performed in previous analyses (see Chapter 2 for details), each individual was identified as *P. bocagei, P. carbonelli* or hybrid. The majority of the field identifications were congruent with the genetic assignments (Figure 8 A vs. B). Once again *P. bocagei* is located in the North and *P. carbonelli* in the South, contacting at a restricted zone at the center, which extends across approximately 500 meters. The majority of hybrid individuals were located at the contact zone, although three individuals recovered as hybrids (NEWHYBRIDS results, see Chapter 2) were sampled at the south of this area, two of which were putative backcrosses resulting from the mating of a first generation hybrid with *P. carbonelli* and another being an F2.

3.3.2. Interpolations

The interpolations performed with the genetic data (proportion of genome assignment and genetic distances) returned very similar results (Figure 9 – A and C). Both show genetic traits associated with *P. bocagei* restricted to the northern part of the study area and traits associated with *P. carbonelli* restricted to the south. At the center of the sampled area, where the contact zone is located, we observe a higher density of genetic traits associated with *P. carbonelli* individuals, but with a high density of *P. bocagei* traits localized across a restricted horizontal stripe. No clear trend is detected for hybrid individuals, showing their spread distribution through the area and the low number of admixed individuals. Both analyses show a really steep transition between the two species located at around 600 meters from the northernmost sample, coincident with the area where most hybrids were located.

Interpolation maps of pholidotic traits show a more patchy distribution of morphological variation across the landscape. However, morphologies associated with *P. bocagei* are still more localized at the north of the study area. Contrary to the previous interpolations based on genetic data, these results show, at the center of the area, the presence of morphologies associated to both parental forms with a slightly higher density of *P. carbonelli* morphologies and at the South both species in a similar ratio (Figure 9 – B). The interpolation based on morphological distances show similar results but with a slightly higher dominance of *P. carbonelli* when compared with the morphologically based PP. Here, *P. bocagei* is the most frequent species at the

northern location of the study area but not as pronounced as in previous analyses (Figure 9 – D).

Some strong differences exist between the distributions of genetic and morphological variability across the landscape. First, the genetic distribution suggests a steep clinal change in genetic traits associated to the two parental forms, whereas morphological traits present a mosaic distribution. At the same time, genetically-based interpolations present a higher density of genetic traits associated with *P. carbonelli* population while morphological traits associated with the two parental forms are found at a similar ratio.

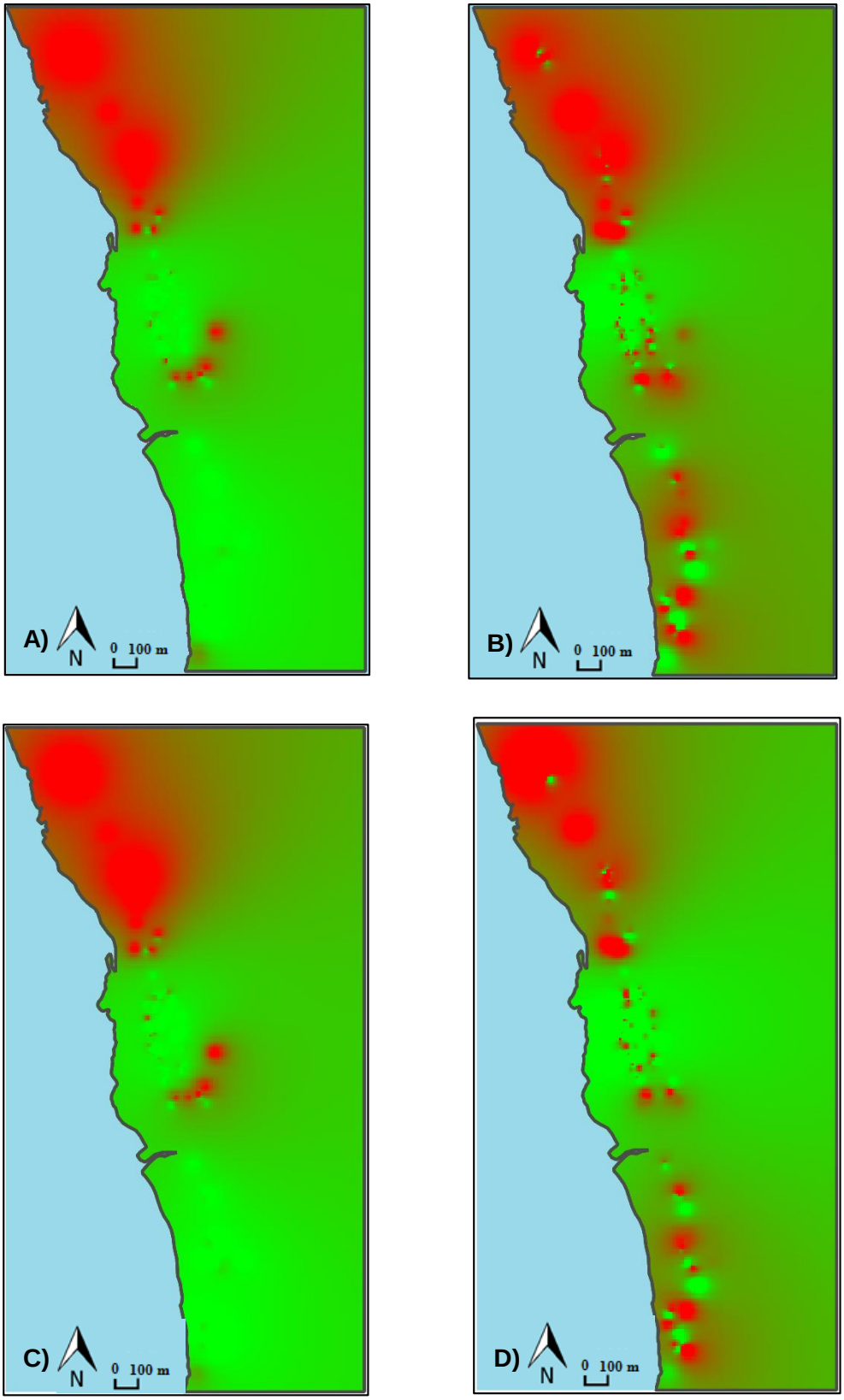


Figure 9 - Inverse distance weighted interpolations. A) Interpolation performed with the proportion of genome assignment estimated by STRUCTURE software; B) interpolation performed with the PP calculated using the morphology-based discriminant analysis; C) interpolation performed with the genetic distances between individuals calculated with microsatellite scores; D) interpolation performed with the Euclidian distances between individuals calculated with morphological characters.

3.3.3. Clines

The variation of best-fitting model was small, having the most common model (chosen for 12/16 loci and for the STRUCTURE assignment cline) a pMin and pMax estimated and fitting tail not estimated (none fitted). For the loci Ph22 and Ph62 the model that best explains the data used the observed values of pMin and pMax and has only one exponential tail at the right side. The model chosen for locus Ph62 used the observed values of pMin and pMax and has mirror tails. Only one microsatellite marker (Ph128) was better explained by the null model that assumes no variation across the landscape. The best-fit model for the mtDNA cline fixed pMin and pMax at 0 and 1 respectively and has just one exponential tail on the right side (Table S3, in Supplementary material).

Inferred clines did not show great differences regarding their centers location with a mean of 649.9m, standard deviation (SD) of 45.7m, ranging from 582.8m to 812.4m. In contrast, cline widths were highly variable (mean = 114.1 ± SD 146.4, range 0.009-591.4m). The cline performed with individual assignments had the center located at 642.9m from the northernmost sampled individual and a width of 124.8m. Mitochondrial DNA cline was centered at 646.6m and had a width of 96.2m.

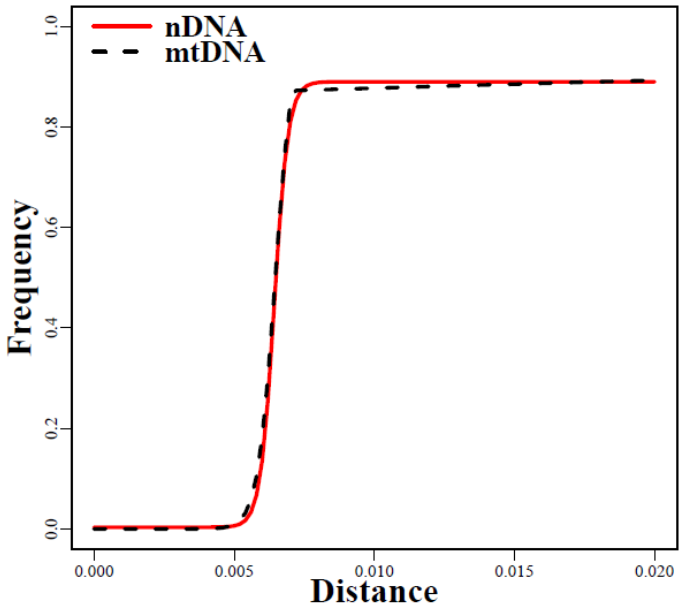


Figure 10 - Combined maximum likelihood geographic clines across the study area for the proportion of genome assignment estimated by STRUCTURE software using 16 microsatellite loci (nDNA) and species classifications using mtDNA (mtDNA). X axis representing the frequency of the genetic marker(s) and y axis representing distance (Kmx10).

The intervals of center values overlap suggesting coincident clines for microsatellite loci, proportion of genome assignment and mtDNA. Despite the variation of cline width all confidence intervals show substantial overlap, thus preventing us to reject differences in this parameter for the distinct analyses.

All clines showed a steep transition along the transect located approximately at the same distance (approximately 600m) from the northernmost sampled individual (Figures 10 and 11).

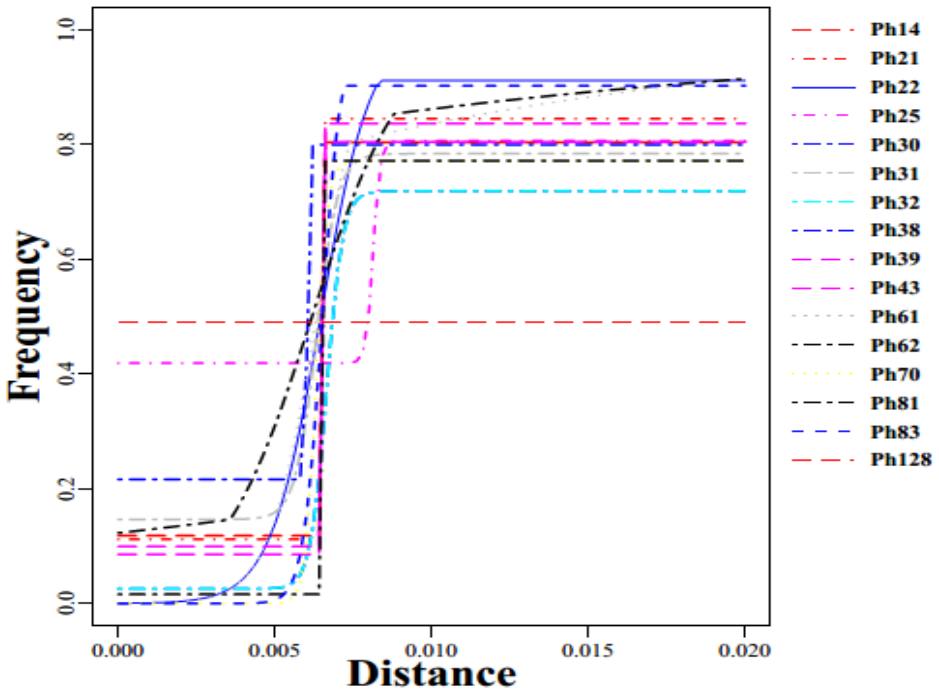


Figure 11 - Combined maximum likelihood geographic clines across the study area for 16 microsatellite loci analysed. X axis representing the frequency of the genetic marker(s) and y axis representing distance (Kmx10).

3.4. Discussion

The results of the present study highlight the importance of spatially explicit frameworks for the study of hybrid zones. These provide useful tools to investigate geographical patterns of morphological and genetic variation, to analyze the dynamics of gene flow and gain insights to the ecological factors shaping species distributions in contact zones (Kidd & Ritchie 2000; Martínez-Freiría *et al.* 2008).

The geographic nature of the hybrid zone: clinal or patchy?

Despite the small size of the study area (only 2.3Km long) our fine-scale spatial analysis reveals a steep clinal change in genetic traits with a very narrow hybrid zone, and with the majority of hybrid individuals being restricted to the area where the two parental species are found in syntopy. Although two of the individuals identified in previous analyses (Chapter 2) as backcrosses with *P. carbonelli* and one identified as F2 hybrid were detected at the south of the sympatric area their location may be easily explained by hybrid dispersal. In fact, a mosaic structured or geographically unstructured hybrid zone were probable at the geographic scale of our study, which is close to the scale of *Podarcis* dispersal (Vignoli *et al.* 2012). Although genetic interpolations provide some support to a patchy distribution, the results of clinal analyses strongly support the existence of a single and steep clinal change across this area. Interpolations of morphological traits also suggest a patchy distribution across the landscape, more congruent with a mosaic hybrid zone. However, these results should be taken with caution, as morphological traits may be influenced by numerous factors other than those driving hybridization and are known to vary extensively in most *Podarcis* species. Furthermore, only three pholidotic continuous characters were use in this analysis and these show extensive overlap between these species (Chapter 2). Such overlap is usual even among unrelated *Podarcis* species, and is probably associated with the high intraespecific variability, typical of the genus *Podarcis* (Kaliontzopoulou *et al.* 2005, 2012b). Other characters that are highly variable across samples such as biometric (Kaliontzopoulou *et al.* 2012b) or coloration characters (Sá-Sousa & Almeida 2000; Sá-Sousa & Harris 2002; Kaliontzopoulou *et al.* 2005, 2012b) and additional clinal analyses performed with morphological data could improve inference regarding spatial distribution of morphological phenotypes.

In summary, whether this hybrid zone fits a clinal or a mosaic pattern is hard to evaluate, since there are evidences favoring both types of genotype and phenotype distribution. It may well be that the highly clinal pattern estimated from cline analyses is a consequence of the rather simplistic cline models that are necessarily fit to the data.

Are the steep clines observed at a microscale caused by a barrier to dispersal?

The steep genetic gradient, the high congruence in cline location and cline width and the restricted location of the hybrids suggests the presence of strong barriers

to gene flow and/or a recent contact between the two species. Clinal analyses show that the hybrid zone, as determined by the genetic markers, is centered at approximately 600m from the northernmost extreme of the study area. At 600m from the northernmost extreme of the sampling area we found a narrow water stream that runs from the fields located closer to the urban area to the sea (Figure 12), dividing our study area. This water stream can be working as a geographic barrier, limiting the dispersal of individuals of *P. bocagei* to the South and of *P. carbonelli* to the North as we found a steep transition between the species at this point. In fact, we were not able to sample a single *P. carbonelli* individual at the North of this stream. It is possible that *P. bocagei* located at the South is isolated from the rest of the population. However, dispersal is still possible as we found *P. bocagei* across all the sympatric area and with a higher density at its southern part (Figure 9 – A and C). Sidewalks can facilitate dispersal across the stream, as it was shown that *Podarcis* could pass along woodpasses (Carretero, personal observation). However, this stream could be more effective as a barrier to dispersal in *P. carbonelli* than in *P. bocagei*, which might partially explain the observed pattern. More detailed analyses (e.g. inferring the relatedness between *P. bocagei* separated by this putative barrier) can provide additional information about the strength of the stream as barrier to dispersal.

The nature of selection against hybrids: endogenous or exogenous?

The location of the center of the hybrid zone at this stream suggests this hybrid zone to be a tension zone, as these hybrid zones tend to move towards and become trapped by barriers to dispersal (Barton & Hewitt, 1985). Tension zones are traditionally viewed as being mostly maintained by endogenous selection (Barton & Hewitt 1985; Kruuk *et al.* 1999) although both types of selection can be present. The present study was performed at a really small geographic scale, with no obvious differences in main environmental features such as altitude, temperature or habitat. However, in the sympatric area, *P. bocagei* is found mostly in the lateral sidewalk of the beach (see Figure 9- A and C), probably due to the adaptation of this species to anthropogenic structures (Galán 2009). In fact, although both species can use anthropogenic structures, *P. bocagei* seems to perform better in urban environments (Galán 2009). Furthermore, *P. bocagei* shows a tendency to occupy areas with higher vegetation and *P. carbonelli* tends to prefer cleaner and more open areas with small vegetation (Carretero *et al.* 2002). At the South of the stream, dune area is wider and with more vegetation and higher environmental heterogeneity at the micro-geographical scale,

what could allow for some ecological displacement between the two species. Also, some intrinsic barriers to gene flow have already been detected between these two species. In fact, male chemical recognition by conspecific females has been demonstrated for both species (Barbosa *et al.* 2005) suggesting assortative mating, although the lack of chemical recognition by females (Barbosa *et al.* 2005) and the hybridization in captivity (Galán 2002) suggest that that particular reproductive barrier is not completely effective. Despite important for evolutionary studies, it is not possible to distinguish the effects of exogenous and endogenous selection using clinal analyses as both types of selection result in similar shaped clines but are important in evolutionary studies (Barton & Hewitt 1985; Kruuk *et al.* 1999). More studies are required to distinguish the roles of intrinsic and extrinsic factors limiting gene flow between these two species. These studies can include captivity breeding, genomic clinal analyses and studies of habitat use by both species.

Figure 12 - Photograph representing the centre of the cline (approximately 600m from the northernmost sampled individual. In the image it is visible the presence of a river stream indicated by a black arrow. Source: "Cline centre of the study area". 41º01´52.66´´N and 8º38´43.86´´E. **Google Earth**. June 26, 2007. November 1, 2014.

Concluding remarks

Previous analyses investigating the genetic and morphological composition of the hybrid zone as well as the existence of developmental instability were not able to find the main reproductive barrier maintaining differentiation in the area. This spatial analysis at the individual level in the contact area, complements previous analyses

suggesting the existence of strong barriers to gene flow between these two lizard species responsible for the bimodal distribution of genotypes, identifying a landscape feature that may be restricting introgression between the two forms, by limiting their contact. This was only possible due to great spatial resolution of this study, closer to the spatial scale of the organism (Tarroso *et al.* 2014).

4. Final remarks

4.1. General discussion

The major aim of the present study was to characterise the hybrid zone between *P. bocagei* and *P. carbonelli* in order to better understand the extent and consequences of hybridisation in their integrity and the barriers limiting gene flow between these species. In order to accomplish our objectives we analysed the genetic and morphological composition of the hybrid zone and the spatial distribution of different characters across the area.

In the present study we used one mtDNA molecular marker and 16 microsatellite markers to distinguish the two species and analyse the patterns of genetic admixture. Although a higher resolution, namely in terms of number and quality of markers, would probably allow to better characterise the admixture process, the simulations performed showed that the genetic markers used were sufficiently informative for the purposes of this study.

Our genetic analysis detected a strong predominance of pure individuals in the contact zone, confirming the bimodality of the hybrid zone reported in a previous study of the area (Pinho *et al.* 2009). Therefore, despite the presence of gene flow and introgression between *P. bocagei* and *P. carbonelli*, the results obtained indicate strong barriers to gene flow in the area suggesting an almost complete reproductive isolation. Hence, despite the existence of hybridization and introgression, it is clear that the integrity of both species has been maintained during the time interval between this study and the previous examination of the same hybrid zone (first sampled more than 10 years ago).

This work brings new evidence for testing of Haldane´s rule mostly due to the finding of F1 hybrids between the species, which were not detected in the previous study of this contact zone. Specifically, we found no evidence for female unviability, but signs of at least partial female sterility were detected, as backcrossed individuals with pure species only carried the maternal lineage of the pure parent. This female sterility could be influencing the low number of hybrids detected.

The investigation of the direction of hybridization as assessed through the examination of the distribution of mtDNA haplotypes and their comparison to nuclear data provides more insights to the mechanisms shaping the genetic characteristics of this hybrid zone. Overall, 67% of hybrid individuals (including all F1 hybrids) carried the maternal lineage of *P. bocagei*, which indicates a higher rate of hybridization between

females of that species with males of *P. carbonelli* in interspecific matings and with hybrid males. The hypotheses concerning the direction of hybridization would need to be explicitly tested, but the pattern observed here suggests asymmetric crossings between males and females of both species. This pattern could be related with several factors, such as the existence of reproductive, genetic and mechanical incompatibilities. Further, assortative mating or asymmetric viability of hybrids could contribute to the patterns observed. The present work provides interesting hypothesis that deserve further research.

A parallel morphological screening of the contact zone between the two *Podarcis* species can help us to understand the cohesiveness of each species and investigate the influence of gene flow on the morphological characteristics of hybrid individuals. We detect a higher similarity between both species in the sympatric area when compared with the allopatric populations. This could be an effect of long-term hybridization, which causes introgression and influences the morphology of the individuals. However, we need to examine alternative explanations, namely, that this result may also arise due to the influence of environmental factors at a local scale. Therefore, morphological changes can appear due to local adaptation or phenotypic plasticity, not being directly influenced by reproductive isolation, or the lack of it thereof (DeWitt & Scheiner 2004). Simultaneously, no relationship was detected between genetic and morphological variation. It is noteworthy that in the literature, morphological diversification frequently does not match the diversification of molecular markers. This suggests that either the evolutionary tempo of morphological change is slower than that of speciation or that the human capacity to detect morphological differences between species is limited (Fritz *et al.* 2006). Whatever the case, it certainly had an influence on our results supported by good sample sizes.

Morphological traits have been analysed in several studies on the variation of *Podarcis* lizards showing that, despite the high inter- and intraspecific variability, it is possible to differentiate the lineages when analyses are performed based on many morphological characters (Kaliontzopoulou *et al.* 2005, 2011, 2012b). In our analyses hybrid individuals do not differ in morphology from the parental forms while no morphologically intermediate forms were identified. No relation was detected between genetic and morphological assignments of hybrid individuals. Therefore, morphological traits do not follow genetic admixture patterns. Note, however, that these statistical results are based on only three pholidotic characters. Combining pholidotic characters with linear measurements or using geometric morphometrics could enhance the detection of intermediate morphologies and reveal association between genetic and morphological patterns of variability, as these methods are known to provide a high

discrimination power and resolution (Kaliontzopoulou 2011; Kaliontzopoulou *et al.* 2012b). Likewise, coloration patterns, not analysed here due to logistic constraints, may be also a powerful diagnostic tool to distinguish between species and examine whether hybrids exhibit signs of intermediate morphologies, as they are already used in field identifications (Sá-Sousa & Almeida 2000; Sá-Sousa & Harris 2002; Kaliontzopoulou *et al.* 2005, 2012b). In fact, colour characters are important in intraspecific communication, and they are therefore potentially involved in species recognition and reproductive isolation (Olsson 1994; Vercken *et al.* 2007; Bajer *et al.* 2010; Perez i de Lanuza *et al.* 2012). It is evident that the analysis of coloration could bring valuable evidence to study barriers to gene flow in hybrid zones, either in lizards or in other visual animals.

Morphological traits were also used in order to investigate the hypothesis of higher developmental instability on hybrids when compared to parental forms. As *Podarcis* are bilateral organisms and many anatomical characters fulfil the criteria for fluctuating asymmetry (FA), this trait could be used to infer the magnitude of developmental instability across the hybrid zone. In fact, FA was already investigated in lizards to understand the influence of environmental stress caused by anthropogenic activities (Lazić *et al.* 2013), or to infer competition (Carretero *et al.* 2003). However, so far, this is the first time FA analysis is applied to hybrid zones in *Podarcis* or other lizards. Our results at the individual level do not indicate differences in developmental instability on hybrids compared with parental forms, giving no indications of lower fitness in hybrids. Nevertheless, since FA could not be analysed at the population level due to the low number of hybrids, this study is only a first approach to the analysis of FA variation in hybrid lacertids.

All the previous results suggest strong barriers to gene flow in the hybrid zone. However, putting these results in a specific spatial framework allowed us to study the spatial dynamics of different genetic and morphological traits across the area and enhance our understanding of the underlying biological mechanisms limiting gene flow between species.

The cline analyses uncover a steep transition across the study area, concordant between distinct genetic markers and coincident with the location of water stream. These results suggest this water stream to be acting as a barrier to dispersal that limits the contact between the two species. However, *P. bocagei* individuals were detected to the south of this stream, showing the permeability of this geographic barrier. The strong bimodal distribution observed in the area of syntopy, together with the evidence from limited hybrid female fertility reported in this study, strongly suggest the existence of reproductive barriers, besides the geographic barrier, limiting hybridisation between

these two species. Both endogenous and exogenous barriers may be involved in the reproductive isolation observed between these two species but further studies are required to identify these barriers and understand their strength and possible role.

Interpolation results from morphological and to a less extent genetic traits, suggest a mosaic structure for this hybrid zone. Morphological results should be interpreted with caution, since only three pholidotic characters were used and additional morphological traits would be required to improve our inference. Also, the high intraspecific variation present in these characters is probably influencing this pattern. However, the apparent differences between the results of interpolation and clinal analyses can be explained by the effect of the geographic barrier that limits the introgression in the northern direction creating a very steep cline between the pure and relatively isolated population of *P. bocagei* located at the north of this stream and the real zone of admixture located at the south of this barrier. It is possible that removing the strongly isolated population of *P. bocagei* at the north of the stream would return more congruent results between clinal analysis and interpolation representations.

4.2. Conclusions and future work

4.2.1. Conclusions

The multidisciplinary approach used in the present study provides broader insights on the characteristics of hybridization between two species of wall lizards and on factors shaping this hybrid zone. Overall, we detect evidence supporting the important role of landscape features (such as geographic barriers) and selection against hybrids in the shape and features of this contact zone. The study shows that only by integrating different sources of evidence (in this case genetic, morphology and geographic distribution) is it possible to extract robust conclusions on the reproductive barriers and selective forces acting in hybrids. The combination of genetic, morphological, and spatial data allows the investigation of the nature and influence of reproductive barriers maintaining species integrity.

4.2.2. Future work

The present study may benefit from a more detail investigation of the factors shaping the hybrid zone, which would allow us to obtain a more complete picture of this interesting study area. Therefore several steps could be perform:

I. In a first step we want to increase the number of specimens analysed in order to sample a higher number of hybrid individuals which could improve the power of statistical analysis;

II. From a practical perspective, the analysis of thousands of loci or even of whole genomes are more accessible nowadays (John *et al.* 2004; Morin *et al.* 2004; Middleton *et al.* 2004). Therefore, using more loci would allow us to better characterise the genetic composition of the hybrid zone and to identify genomic regions under selection and/or involved in the reproductive isolation of the species and study its relation with environmental features;

III. As we know that the species considered in this project are able to hybridize in captivity a good approach for the morphological analyses will be perform controlled experiments in the laboratory in order to analyse the morphology of the progeny;

IV. We want to increase the number of morphological characters and use geometric morphometrics to analyse the morphological properties of the hybrids into more detail as well as evaluate levels of asymmetry. Specially, we want to understand if the lack of differences detected between hybrids and parental forms is biased by the number and type of morphological traits used. Through the combination of the analysed traits with new ones we will be able to test with higher confidence if the negative selection of hybrids is related with phenotypic characteristics;

V. On a different perspective we want to perform a relatedness analysis in order to compare the relations between individuals with their spatial distribution in order to test for spatial segregation of families. Also, a parallel study of home ranges is being performed with individuals of the syntopic area which would increase the knowledge on the spatial patterns detected on the hybrid zone;

VI. In order to identify the position of offspring in relation to their parents, paternity analysis can be perform to provide an indirect measure of dispersal, which

would allow a more detailed analysis of the spatial structure of hybrids compared with parental forms.

VII. The application of this type of multidisciplinary approaches in other *Podarcis* hybrid zones would increase the knowledge about the pattern of hybridization in the genus, allowing the study of the causes and consequences of speciation in the taxa. It is also important to refer that this project is connected with a broader study of contact zones between different pair of *Podarcis* species in the Iberian Peninsula.

5. References

Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–46.

Adams DC, Berns CM, Kozak KH, Wiens JJ (2009) Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society of London B*, **276**, 2729–38.

Agostini C, Albaladejo RG, Aparicio A *et al.* (2013) Permanent genetic resources added to molecular ecology resources database 1 April 2013-31 May 2013. *Molecular Ecology Resources*, **13**, 966–8.

Alexandrino J, Baird S, Lawson L (2005) Strong selection against hybrids at a hybrid zonein the ensatina ring species complex and its evoutionary implications. *Evolution,* **59.6**, 1334-1347.

Alibert P, Renaud S, Dod B, Bonhomme F, Auffray JC (1994) Fluctuating asymmetry in the *Mus musculus* hybrid zone: a heterotic effect in disrupted co-adapted genomes. *Proceedings of the Royal Society of London B*, **258**, 53–9.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution*, **16**, 613–622.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217–1229.

Arnold E (1973) Relationships of the Palaearctic lizards assigned to the genera Lacerta, *Algyroides* and *Psammodromus* (Reptilia: Lacertidae), London: British Museum (Natural History).

Arnold E (1987) Resource partition among lacertid lizards in southern Europe. *Journal of Zoology,* **1.4**, 739-782

Arnold E (1990) Why do morphological phylogenies vary in quality? An investigation based on the comparative history of lizard clades. *Proceedings of the Royal Society of London B, **240.1297**, 135-172.*

Arnold M (1992) Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics*, **23**, 237–261.

Arnold E, Burton J, Ovenden D (1978) Field guide to the reptiles and amphibians of Britain and Europe.

Arntzen J, Sá-Sousa P (2007) Morphological and genetical differentiation of lizards (*Podarcis bocagei* and *P. hispanica*) in the Ria de Arosa Archipelago (Galicia, Spain) resulting from vicariance. *Biogeography, time and place: distributions, barriers, and islands*. Springer Netherlands. 365-401.

Avise J (1994) Molecular Markers, Natural History and Evolution. *The Auk* **121.4**, 1298-1299.

Baack EJ, Rieseberg LH (2007) A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development*, **17**, 513–8.

Bajer K, Molnár O, Török J, Herczeg G (2010) Female European green lizards (Lacerta viridis) prefer males with high ultraviolet throat reflectance. *Behavioral Ecology and Sociobiology*, **64**, 2007–2014.

Barbosa D, Desfilis E, Carretero M, Font E (2005) Chemical stimuli mediate species recognition in Podarcis wall lizards. *Amphibia-Reptilia*, **26**, 257–263.

Barbosa D, Font E, Desfilis E, Carretero MA (2006) Chemically mediated species recognition in closely related *Podarcis* wall lizards. *Journal of Chemical Ecology*, **32**, 1587–98.

Barilani M, Bernard-Laurent A, Mucci N *et al.* (2007) Hybridisation with introduced chukars (*Alectoris chukar*) threatens the gene pool integrity of native rock (*A. graeca*) and red-legged (*A. rufa*) partridge populations. *Biological Conservation*, **137**, 57–69.

Barilani M, Deregnaucourt S, Gallego S *et al.* (2005) Detecting hybridization in wild (*Coturnix coturnix*) and domesticated (*Coturnix japonica*) quail populations. *Biological Conservation*, **126**, 445–455.

Barton NH (1979) Gene flow past a cline. *Heredity*, **43**, 333–339.

Barton N, Bengtsson BO (1986) The barrier to genetic exchange between hybridising populations. *Heredity*, **57**, 357–76.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Barton N, Hewitt G (1989) Adaptation, speciation and hybrid zones. *Nature,* **341.6242**, 497-503.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *Laboratoire Génome, Populations, Interactions, CNRS UMR* 5000, 1996-2004.

Benson JF, Patterson BR, Wheeldon TJ (2012) Spatial genetic and morphologic structure of wolves and coyotes in relation to environmental heterogeneity in a Canis hybrid zone. *Molecular Ecology*, **21**, 5934–54.

Braña F (1996) Sexual dimorphism in lacertid lizards: male head increase vs female abdomen increase? *Oikos*, **75**, 511–523.

Britch SC, Cain ML, Howard DJ (2001) Spatio-temporal dynamics of the *Allonemobius fasciatus* - A. socius mosaic hybrid zone: a 14-year perspective. *Molecular Ecology*, **10**, 627–638.

Bronson C, Grubb T, Braun M (2003) A test of the endogenous and exogenous selection hypotheses for the maintenance of a narrow avian hybrid zone. *Evolution*, **57.3**, 630-637

Brumfield RT, Jernigan RW, McDonald DB, Braun MJ (2001) Evolutionary implications of divergent clines in an avian (Manacus: Aves) hybrid zone. *Evolution*, **55**, 2070–2087.

Butlin R, Debelle A, Kerth C *et al.* (2012) What do we need to know about speciation? *Trends in Ecology & Evolution*, **27**, 27–39.

Camargo A, Sinervo B, Sites J (2010) Lizards as model organisms for linking phylogeographic and speciation studies. *Molecular Ecology*, **19.16**, 3250-3270.

Capula M (1993) Natural hybridization in *Podarcis* sicula and *P. wagleriana* (Reptilia: Lacertidae). *Biochemical Systematics and Ecology*, **21.3**, 373-380.

Capula M (2002) Genetic evidence of natural hybridization between *Podarcis sicula* and *Podarcis tiliguerta* (Reptilia: Lacertidae). *Amphibia-Reptilia*, **23**, 313–321.

Capula M, Ceccarelli A (2003) Distribution of genetic variation and taxonomy of insular and mainland populations of the Italian wall lizard, *Podarcis sicula*. *Amphibia Reptilia*, **24.4**, 483-496

Capula M, Luiselli L, Bologna M, Ceccarelli A (2002) The decline of the Aeolian wall lizard, *Podarcis raffonei*: causes and conservation proposals. *Oryx*, **36.01,** 66-72.

Carretero M a (2008) An integrated assessment of a group with complex systematics: the Iberomaghrebian lizard genus *Podarcis* (Squamata, Lacertidae). *Integrative Zoology*, **3**, 247–66.

Carretero M, Barbosa D, Ribeiro R (2003) Assessing competition by means of fluctuating asymmetry in *Podarcis*. *12th Ordinary General Meeting, Societas Europaea Herpetologica. St. Petersburg, Russia, 12-16 August 2003*.

Carretero MA, Harris DJ, Ribeiro R, Barbosa D, Sá-Sousa P (2006) Spermatogenesis in two Iberian *Podarcis* lizards: relationships with male traits. *Animal Biology*, **56**, 1–12.

Carretero MA, Llorente GA (1993) Morfometría en una comunidad de lacértidos mediterráneos y su relación con la ecología . *Historia Animalium*, **2**, 77–99.

Carretero MA, Sa-Sousa P, Barbosa D, Harris DJ, Pinho C (2002) Sintopía estricta entre *Podarcis bocagei* y *P. carbonelli*. *Boletín de la Asociación Herpetológica Española*, **13**, 20–24.

Cockerham CC, Weir BS (1984) Covariances of relatives stemming from a population undergoing mixed self and random mating. *Biometrics*, **40**, 157–64.

Cole C (1966) Femoral glands in lizards: a review. *Herpetologica*, 199-206.

Corander J, Waldmann P, Sillanpaa MJ (2003) Bayesian Analysis of Genetic Differentiation Between Populations. *Genetics*, **163**, 367–374.

Cornuet J-M, Piry S, Luikart G, Estoup A, Solignac M (1999) New Methods Employing Multilocus Genotypes to Select or Exclude Populations as Origins of Individuals. *Genetics*, **153**, 1989–2000.

Corti C, Masseti M, Delfino M, Pérez-Mellado V (1999) Man and herpetofauna of the Mediterranean islands. *Revista Española de Herpetología*, **13**, 83-100.

Coyne JA, Orr H (2004) *Speciation*. Sinauer Associates, Sunderland, MA.

Derryberry EP, Derryberry GE, Maley JM, Brumfield RT (2014) HZAR: hybrid zone analysis using an R software package. *Molecular Ecology Resources*, **14**, 652–63.

DeWitt T, Scheiner S (2004) Phenotypic variation from single genotypes. *Phenotypic plasticity: functional and conceptual approaches. Oxford, Oxford University Press, 2004.*

Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. *American Naturalist*. 312-321.

Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. *Nature*. **421.6920**, 259-264.

Dow B, Ashley M (1998) High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. *Journal of Heredity.* **89. 1**, 62-70.

Dowling T, Secor C (1997) The role of hybridization and introgression in the diversification of animals. *Annual review of Ecology and Systematics*. 593-619.

Eeva T, Tanhuanpää S (2000) Biomarkers and fluctuating asymmetry as indicators of pollution-induced stress in two hole-nesting passerines. *Functional Ecology. **14.2,** 235-243.*

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Review Genetics*, **5**, 435–445.

Estoup A, Garnery L, Solignac M, Cornuet J (1995) Microsatellite variation in honey bee (*Apis mellifera L.*) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics*. **140.2**, 679-695.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online.* **1**, 47.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics,* **131.2**, 479–491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–87.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology notes*, **7**, 574–578.

Ferrand de Almeida N, Ferrand de Almeida, P. Gonçalves H, Sequeira F, Teixeira J, Ferrand de Almeida F (2001) Guias FAPAS anfíbios e répteis de Portugal. *FAPAS, Oporto*.

Fitze PS, Galliard JF Le, Federici P, Richard M, Clobert J (2005) Conflict over multiple-partner mating between males and females of the polygynandrous commom lizards. *Evolution*, **59**, 2451–2459.

Fong JJ, Chen T-H (2010) DNA evidence for the hybridization of wild turtles in Taiwan: possible genetic pollution from trade animals. *Conservation Genetics*, **11**, 2061–2066.

Fritz U, d'Angelo S, Pennisi M, Valvo M (2006) Variation of Sicilian pond turtles, *Emys trinacris*-What makes a species cryptic? *Amphibia-Reptilia*. **27.4**, 513-529.

Galán P (1997) Reproductive ecology of the lacertid lizard *Podarcis bocagei*. *Ecography*. **20.2**, 197-209

Galán P (1999) Demography and population dynamics of the lacertid lizard *Podarcis bocagei* in north-west Spain. *Journal of Zoology*, **249**, 203–219.

Galán P (2002) Hibridación en laboratorio de *Podarcis bocagei* y *Podarcis carbonelli*. *Boletín de la Asociación Herpetológica Española.* **13**, 20-24.

Galán P (2009) Lagartija de Bocage - *Podarcis bocagei*. *Enciclopédia Virtual de los Vertebrados Españoles.*

García-Muñoz E, Sillero N (2010) Two new types of noose for capturing herps. *Acta Herpetologica*. **5.2**, 259-264.

Gay L, Crochet P-A, Bell D a, Lenormand T (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution*. **62**, 2789–806.

Geniez P, Sá-Sousa P, Guillaume CP, Cluchier A, Crochet P-A (2014) Systematics of the *Podarcis hispanicus* complex (Sauria, Lacertidae) III: valid nomina of the western and central Iberian forms. *Zootaxa*, **3794**, 1–51.

Gifford M (2008) Divergent character clines across a recent secondary contact zone in a Hispaniolan lizard. *Journal of Zoology*, **274.3**, 292-300.

Glossip D, Losos J (1997) Ecological correlates of number of subdigital lamellae in anoles. *Herpetologica*, 192-199.

Gompert Z, Parchman TL, Buerkle CA (2012) Genomics of isolation in hybrids. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **367**, 439–50.

Grosjean P (2014) SciViews: A GUI API for R. *UMONS, Mons, Belgium*.

Guichoux E, Lagache L, Wagner S *et al.* (2011) Current trends in microsatellite genotyping. *Molecular Ecology Resources*, **11**, 591–611.

Haldane JBS (1922) Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, **12**, 101–109.

Harris D, Arnold E (1999) Relationships of wall lizards, *Podarcis* (Reptilia: Lacertidae) based on mitochondrial DNA sequences. *Copeia*, 749-754.

Harris D, Pinho C, Carretero M, Corti C, Bohme W (2005) Determination of genetic diversity within the insular lizard Podarcis tiliguerta using mtDNA sequence data, with a reassessment of the phylogeny of *Podarcis*. *Amphibia Reptilia*, **26.3**, 401.

Harris DJ, Sa-Sousa P (2002) Molecular phylogenetics of Iberian wall lizards (*Podarcis*): Is *Podarcis hispanica* a species complex? *Molecular Phylogenetics and Evolution*, **23**, 75–81.

Harris D, Sá-Sousa P (2001) Species distinction and relationships of the western Iberian *Podarcis* lizards (Reptilia, Lacertidae) based on morphology and mitochondrial DNA sequences. *Herpetological Journal*, **11.4**, 129-136.

Harrison R (1990) Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, **7**, 69-128.

Harrison R (1993) Hybrids and hybrid zones: historical perspective. *Hybrid Zones and the Evolutionary Process*. Oxford University Press.

Hewitt GM (1988a) Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution (Personal edition)*, **3**, 158–167.

Hewitt GM (1988b) Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution (Personal edition)*, **3**, 158–167.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Holderegger R, Wagner H (2006) A brief guide to landscape genetics. *Landscape Ecology*, **21.6**, 793-796.

Holderegger R, Wagner H (2008) Landscape genetics. *Bioscience*, **58**, 199.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–32.

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, **15**, 250–255.

John S, Shephard N, Liu G *et al.* (2004) Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. *The American Journal of Human Genetics*. **75.1**, 54–64.

Johnson P, Haydon D (2007) Maximum-likelihood estimation of allelic dropout and false allele error rates from microsatellite genotypes in the absence of reference data. *Genetics*, **175.2**, 827-842.

Kaliontzopoulou A (2004) Efecto de la simpatría en la morfología de dos especies del género *Podarcis* en Portugal. *DEA Thesis, University of Barcelona*.

Kaliontzopoulou A (2011) Geometric morphometrics in herpetology: modern tools for enhancing the study of morphological variation in amphibians and reptiles.*Basic and Applied Herpetology.* **25**, 5-32

Kaliontzopoulou A, Adams DC, van der Meijden A, Perera A, Carretero MA (2012a) Relationships between head morphology, bite performance and ecology in two species of *Podarcis* wall lizards. *Evolutionary Ecology*, **26**, 825–845.

Kaliontzopoulou A, Carretero MA, Llorente GA (2005) Differences in the pholidotic patterns of *Podarcis bocagei* and *Podarcis carbonelli* and possible implications for species determination. *Revista Española de Herpetología*, **19**, 71–86.

Kaliontzopoulou A, Carretero MA, Llorente GA (2007) Multivariate and geometric morphometrics in the analysis of sexual dimorphism variation in *Podarcis* lizards. *Journal of Morphology*, **268**, 152–65.

Kaliontzopoulou A, Carretero MA, Llorente GA (2008) Head shape allometry and proximate causes of head sexual dimorphism in *Podarcis* lizards: joining linear and geometric morphometrics. *Biological Journal of the Linnean Society*, **93**, 111–124.

Kaliontzopoulou A, Carretero MA, Llorente GA (2010) Intraspecific ecomorphological variation: linear and geometric morphometrics reveal habitat-related patterns within *Podarcis bocagei* wall lizards. *Journal of Evolutionary Biology*, **23**, 1234–44.

Kaliontzopoulou A, Carretero MA, Llorente GA (2012b) Morphology of the *Podarcis* wall lizards (Squamata: Lacertidae) from the Iberian Peninsula and North Africa: patterns of variation in a putative cryptic species complex. *Zoological Journal of the Linnean Society*, **164**, 173–193.

Kaliontzopoulou A, Pinho C, Harris DJ, Carretero MA (2011) When cryptic diversity blurs the picture: a cautionary tale from Iberian and North African *Podarcis* wall lizards. *Biological Journal of the Linnean Society*, **103**, 779–800.

Kawakami T, Butlin RK (2001) Hybrid Zones. In: *eLS* . John Wiley & Sons, Ltd.

Kelkar YD, Strubczewski N, Hile SE *et al.* (2010) What is a microsatellite: a computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC repeats. *Genome Biology and Evolution*, **2**, 620–35.

Kidd DM, Ritchie MG (2000) Inferring the patterns and causes of geographic variation in Ephippiger ephippiger (Orthoptera, Tettigoniidae) using geographical information systems (GIS). *Biological Journal of the Linnean Society*, **71**, 269–295.

Kingston SE, Jernigan RW, Fagan WF, Braun D, Braun MJ (2012) Genomic variation in cline shape across a hybrid zone. *Ecology and Evolution*, **2**, 2737–48.

Kisel Y, Barraclough TG (2010) Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, **175**, 316–34.

Klingenberg CP (2002) Morphometrics and the role of the phenotype in studies of the evolution of developmental mechanisms. *Gene*, **287**, 3–10.

Knight M, Turner G (1998) Microsatellite paternity analysis on captive Lake Malawi cichlids supports reproductive isolation by direct mate choice. *Molecular Ecology, **7.11**, 1605-1610.*

Kocher TD, Thomas WK, Meyer a *et al.* (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 6196–200.

Kruuk LEB, Baird SJE, Gale KS, Barton NH (1999) A Comparison of Multilocus Clines Maintained by Environmental Adaptation or by Selection Against Hybrids. *Genetics*, **153**, 1959–1971.

Lazić MM, Kaliontzopoulou A, Carretero MA, Crnobrnja-Isailović J (2013) Lizards from urban areas are more asymmetric: using fluctuating asymmetry to evaluate environmental disturbance. *PloS ONE*, **8**, e84190.

Leaché AD, Cole CJ (2007) Hybridization between multiple fence lizard lineages in an ecotone: locally discordant variation in mitochondrial DNA, chromosomes, and morphology. *Molecular Ecology*, **16**, 1035–54.

Lewontin R, Birch L (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution*, **20**, 315–336.

Li J, Australia G (2008) *A review of spatial interpolation methods for environmental scientists*. Geoscience Australia, Record 2008/23, 137 pp.

Lima A, Pinho C, Larbes S *et al.* (2009) Relationships of *Podarcis* wall lizards from Algeria based on mtDNA data. *Amphibia-Reptilia*, **30**, 483–492.

Loureiro A, Ferrand de Almeida N, Carretero MA, Paulo OS (eds.) (2008) *Atlas dos Anfíbios e Répteis de Portugal*. Instituto da Conservação da Natureza e da Biodiversidade, Lisbon, 257 pp.

Lu G, Basley DJ, Bernatchez L (2001) Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. *Molecular Ecology*, **10**, 965–85.

Lu GY, Wong DW (2008) An adaptive inverse-distance weighting spatial interpolation technique. *Computers & Geosciences*, **34**, 1044–1055.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, **20**, 229–237.

Mallet J (2007) Hybrid speciation. *Nature*, **446.7133**, 279-283.

Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B, Biological Sciences*, **363**, 2971–2986.

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology & Evolution*, **20**, 136–42.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

Martínez-Freiría F, Santos X, Pleguezuelos JM, Lizana M, Brito JC (2008) Geographical patterns of morphological variation and environmental correlates in contact zones: a multi-scale approach using two Mediterranean vipers (Serpentes). *Journal of Zoological Systematics and Evolutionary Research*, **47**, 357–367.

Mayr E (1942) Systematics and the origin of species, from the viewpoint of a zoologist. *Harvard University Press*

Mayr E (1963) Animal Species and Evolution. *Belknap Press of Harvard University Press.*

Middleton F a, Pato MT, Gentile KL *et al.* (2004) Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: a comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. *American Journal of Human Genetics*, **74**, 886–97.

Møller A (1992) Parasites differentially increase the degree of fluctuating asymmetry in secondary sexual characters. *Journal of Evolutionary Biology*, **5.4**, 691-699.

Moore W, Price J, Harrison R (1993) Nature of selection in the northern flicker hybrid zone and its implications for speciation theory. *Hybrid Zones and the Evolutionary Process, 196-225.*

Morin P a., Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208–216.

Murphy RW, Fu J, Macculloch RD, Darevsky IS, Kupriyanova LA (2000) A fine line between sex and unisexuality: the phylogenetic constraints on parthenogenesis in lacertid lizards. *Zoological Journal of the Linnean Society*, **130**, 527–549.

Nielsen EE, Bach LA, Kotlicki P (2006) Hybridlab (Version 1.0): a Program for Generating Simulated Hybrids From Population Samples. *Molecular Ecology Notes*, **6**, 971–973.

Nosil P (2008) Speciation with gene flow could be common. *Molecular Ecology*, **17.9**, 2103–2106.

Oliverio M, Bologna M, Mariottini P (2000) Molecular biogeography of the Mediterranean lizards *Podarcis* Wagler, 1830 and Teira Gray, 1838 (Reptilia, Lacertidae). *Journal of Biogeography*, **27.6**, 1403-1420.

Olmo E (2005) Chromorep: a reptiles chromosomes database. *Internet reference retrieved from http://193.206* 118, 100.

Olsson M (1994) Nuptial coloration in the sand lizard, Lacerta agilis: an intra-sexually selected cue to lighting ability. *Animal Behaviour*, **48**, 607–613.

Olsson M, Gullberg A, Tegelström H, Madsen T, Shine R (1994) Promiscuous lizards: females have more viable young. *Nature*, **369**, 528.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes*, **4**, 535–538.

Orr H (1995) The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics*, **139.4**, 1805-1813.

Orr HA (1997) Haldane´s rule. *Annual Review of Ecology and Systematics,* 195–218.

Palmer AR (1994) *Fluctuating asymmetry analyses: A primer* (TA Markow, Ed,). Springer Netherlands, Dordrecht.

Palmer AR, Strobeck C (1986) Fluctuating asymmetry - measurement, analysis, patterns. *Annual Review of Ecology and Systematics*, **17**, 391–421.

Palmer AR, Strobeck C (2003) Fluctuating asimmetry analyses revisited. In: Polak M (ed) Developmental Instability (DI): Causes and Consequences. *Oxford University Press*, 279–319.

Pálsson B, Pálsson F, Perlin M (1999) Using quality measures to facilitate allele calling in high-throughput genotyping. *Genome Research,* **9.10**, *1002-1012.*

Peakall R, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England)*, **28**, 2537–9.

Pérez i de Lanuza G, Font E, Carazo P (2012) Color-assortative mating in a color-polymorphic lacertid lizard. *Behavioral Ecology*, **24**, 273–279.

Pérez-Mellado V (1981) Nuevos datos sobre la sistemática y distribución de *Podarcis bocagei* (Seoane, 1884)(Sauria, Lacertidae) en la Península Ibérica. *Amphibia-Reptilia*, **2.3**, 259-265.

Pérez-Mellado V (1982) Algunos datos sobre la reproducción de dos especies de Lacertidae (Sauria, Reptilia) en el Sistema Central. *Boletín de la Real Sociedad Española de Historia Natural (Sección Biologia)* **80.3**, 165-173.

Phillips B, Baird S, Moritz C (2004) When vicars meet: a narrow contact zone between morphologically cryptic phylogeographic lineages of the rainforest skink, *Carlia rubrigularis*. *Evolution*, **58**, 1536–1548.

Piálek J, Barton N (1997) The spread of an advantageous allele across a barrier: the effects of random drift and selection against heterozygotes. *Genetics*, **145.2**, 493-504.

Pinho C (2007) Evolution of wall lizards (*Podarcis* spp.) in the Iberian Peninsula and North Africa (PhD thesis). University of Oporto.

Pinho C, Ferrand N, Harris DJ (2006) Reexamination of the Iberian and North African *Podarcis* (Squamata : Lacertidae) phylogeny based on increased mitochondrial DNA sequencing. *Molecular Phylogenetics and Evolution*, **38**, 266–273.

Pinho C, Harris DJ, Ferrand N (2007a) Comparing patterns of nuclear and mitochondrial divergence in a cryptic species complex: the case of Iberian and North African wall lizards (*Podarcis*, Lacertidae). *Biological Journal of the Linnean Society*, **91**, 121–133.

Pinho C, Harris D, Ferrand N (2007b) Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. *Molecular Ecology*, **16.6**, 1191-1205.

Pinho C, Harris DJ, Ferrand N (2008) Non-equilibrium estimates of gene flow inferred from nuclear genealogies suggest that Iberian and North African wall lizards (*Podarcis* spp.) are an assemblage of incipient species. *BMC Evolutionary Biology*, **8**.

Pinho C, Kaliontzopoulou A, Carretero MA, Harris DJ, Ferrand N (2009) Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *Podarcis carbonelli*: evidence for limited natural hybridization and a bimodal hybrid zone. *Journal of Zoological Systematics and Evolutionary Research*, **47**, 368–377.

Pinho C, Kaliontzopoulou A, Harris DJ, Ferrand N (2011) Recent evolutionary history of the Iberian endemic lizards *Podarcis bocagei* (Seoane, 1884) and *Podarcis carbonelli* Pérez-Mellado, 1981 (Squamata: Lacertidae). *Zoological Journal of the Linnean Society,* **162.1**, 184-200.

Poulakakis N, Lymberakis P (2003) Molecular phylogeny and biogeography of the wall-lizard *Podarcis* erhardii (Squamata: Lacertidae). *Molecular Phylogenetics and Evolution,* **28.1**, 38-46.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

QGIS Development Team (2014) QGIS Geographic Information System.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, **152**, 713–727.

Ross CL, Harrison RG (2002) A fine-scale spatial analysis of the mosaic hybrid zone between gryllus firmus and gryllus pennsylvanicus. *Evolution*, **56**, 2296–2312.

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Sacchi R, Scali S, Pellitteri-Rosa D *et al.* (2010) Photographic identification in reptiles: a matter of scales. *Amphibia-Reptilia*, **31**, 489–502.

Sætre G, Král K, Bures S, Ims R (1999) Dynamics of a clinal hybrid zone and a comparison with island hybrid zones of flycatchers (*Ficedula hypoleuca* and *F. albicollis*). *Journal of Zoology*, **247.01**, 53-64.

Sá-Sousa P (2001) Comparative chorology between *Podarcis bocagei* and *P. carbonelli* (Sauria: Lacertidae) in Portugal. Revista Española de Herpetología 15, 85-97. *Revista Española de Herpetología* , **15**, 85–97.

Sá-Sousa P (2009) Lagartija de Carbonelli - *Podarcis carbonelli*. *Enciclopédia Virtual de los Vertebrados Españoles.*

Sá-Sousa P, Almeida A (2000) Genetic and morphological relationships of the Berlenga wall lizard (*Podarcis bocagei berlengensis*: Lacertidae). *Journal of Zoological Systematics and Evolutionary Research,* **38.2**, 95-102.

Sá-Sousa P, Harris D (2002) *Podarcis carbonelli* Pérez-Mellado, 1981 is a distinct species. *Amphibia Reptilia*, **23.4**, 459-468.

Sá-Sousa P, Pérez-Mellado V, Martínez-Solano I (2009) *Podarcis carbonelli*. The IUCN Red List of Threatened Species. *Version 2014.3*.

Sá-Sousa P, Vicente L, Crespo E (2002) Morphological variability of *Podarcis hispanica* (Sauria: Lacertidae) in Portugal. *Amphibia Reptilia*, **23.1**, 55-70.

Savage A, Hogarth P (1999) An analysis of temperature-induced fluctuating asymmetry in *Asellus aquaticus* (Linn.). *Hydrobiologia*, **411**, 139-143.

Schoville S, Bonin A (2012) Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution, and Systematics,* **43**, 23-43.

Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, **19**, 198–207.

Sequeira F, Alexandrino J, Rocha S, Arntzen JW, Ferrand N (2005) Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander, *Chioglossa lusitanica*. *Molecular Ecology*, **14**, 245–254.

Shurtliff QR, Murphy PJ, Matocq MD (2014) Ecological segregation in a small mammal hybrid zone: habitat-specific mating opportunities and selection against hybrids restrict gene flow on a fine spatial scale. *Evolution.* **68**, 729–42.

Sillero N, Carretero MA (2013) Modelling the past and future distribution of contracting species. The Iberian lizard *Podarcis carbonelli* (Squamata: Lacertidae) as a case study. *Zoologischer Anzeiger - A Journal of Comparative Zoology*, **252**, 289–298.

Smith D (1997) Fluctuating asymmetry in the honey bee, *Apis mellifera*: effects of ploidy and hybridization. *Journal of Evolutionary Biology,* **10.4**, 551–574.

Sobel J, Chen G, Watt L, Schemske D (2010) The biology of speciation. *Evolution*, **64.2**, 295-315.

StatSoft I (2011) STATISTICA, ver. 10.

Stuart-Fox D, Godinho R, de Bellocq JG *et al.* (2009) Variation in Phenotype, Parasite Load and Male Competitive Ability across a Cryptic Hybrid Zone. *Plos ONE*, **4**.

Szymura J, Barton NH (1986) Genetic Analysis of a Hybrid Zone Between the Fire-Bellied Toads , *Bombina bombina* and *B . variegata*, Near Cracow in Southern Poland Published by : Society for the Study of Evolution Stable URL : ht. *Evolution*, **40**, 1141–1159.

Szymura J, Barton N (1991) The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution*, 237-261.

Tarroso P, Pereira RJ, Martínez-Freiría F, Godinho R, Brito JC (2014) Hybridization at an ecotone: ecological and genetic barriers between three Iberian vipers. *Molecular Ecology*, **23**, 1108–23.

Van Tienhoven a. M, Den Hartog JE, Reijns R a., Peddemors VM (2007) A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*. *Journal of Applied Ecology*, **44**, 273–280.

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends in Ecology & Evolution*, **16**, 330–343.

Uetz P, Goll J, Hallermann J (2014) The Reptile Database.

Vähä J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63–72.

Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data, *Molecular Ecology Notes,* **2.3**, 377-379.

Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. *BioTechniques*, **37**, 226–31.

Vercken E, Massot M, Sinervo B, Clobert J (2007) Colour variation and alternative reproductive strategies in females of the common lizard Lacerta vivipara. *Journal of Evolutionary Biology*, **20**, 221–32.

Vignoli L, Vuerich V, Bologna MA (2012) Experimental study of dispersal behaviour in a wall lizard species (*Podarcis sicula*) (Sauria Lacertidae). *Ethology Ecology & Evolution*, **24**, 244–256.

Wiens J (2004) What is speciation and how should we study it? *The American Naturalist*, **163.6**, 914-923.

Willmore K, Young N, Richtsmeier J (2007) Phenotypic variability: its components, measurement and underlying developmental processes. *Evolutionary Biology*, **34.3-4**, 99-120.

Wilsey B, Haukioja E, Koricheva J, Sulkinoja M (1998) Leaf fluctuating asymmetry increases with hybridization and elevation in tree-line birches. *Ecology*, **79.6**, 2092-2099.

Wilson P (1992) On inferring hybridity from morphological intermediacy. *Taxon*, 11-23.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology*, **14**, 851–865.

Wu K, Tanksley S (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. *Molecular and General Genetics MGG*, **241.1-2**, 225-235.

Zhang D, Hewitt G (1996) Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology & Evolution*, **11.6**, 247-251.

6. Supplementary material

Table S1 – Distribution of microsatellite loci in multiplexes and their respective primer tails with fluorescent label, and concentration of primer in each PCR reaction (in uM). All forward primers in amplification reaction were used with a 10x diluted.

Multiplex	Locus	Primer tail	Conc. (µM)
MpA	Ph14	FAM	0.10
	Ph31	VIC	0.10
	Ph32	PET	0.10
	Ph59	NED	0.10
MpB	Ph22	PET	0.10
	Ph35	FAM	0.08
	Ph39	NED	0.06
	Ph170	VIC	0.08
MpC	Ph17	NED	0.20
	Ph21	VIC	0.10
	Ph30	FAM	0.14
	Ph38	PET	0.11
	Ph50	VIC	0.07
MpD	Ph25	VIC	0.08
	Ph83_11	PET	0.20
	Ph142_8	VIC	0.12
MpE	Ph43	FAM	0.30
	Ph70	PET	0.20
	Ph81	FAM	0.14
	Ph128	VIC	0.14
MpF	Ph61	VIC	0.10
	Ph62	PET	0.20

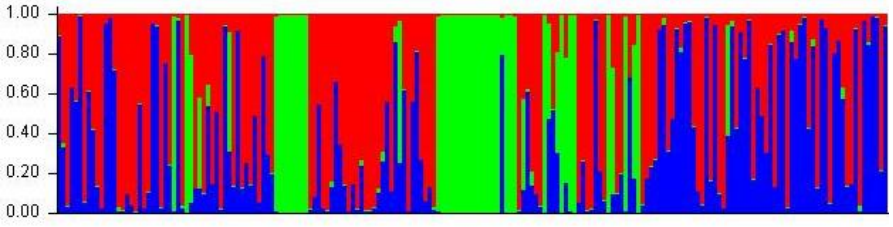


Figure S1 - Model-based multilocus genotype analyses performed in STRUCTURE showing the proportion of the genome of each genotype originating in each of the two species at K=3 (bottom graph). Each individual is represented by a vertical bar divided in two segments, the length of which is proportional to the estimated percentage of the genome of *P. bocagei* (red) or *P. carbonelli* (green).

Table S2 – Maximum likelihood per-allele error rate including allele dropout and false allele of each locus analysed. Analysis perform in PEDANT software using 1000 search steps.

Locus	Allele dropout	False allele
Ph14	0.000	0.000
Ph17	0.025	0.000
Ph21	0.000	0.000
Ph22	0.000	0.000
Ph25	0.000	0.000
Ph30	0.000	0.000
Ph31	0.014	0.000
Ph32	0.000	0.013
Ph35	0.000	0.000
Ph38	0.013	0.000
Ph39	0.000	0.000
Ph43	0.000	0.000
Ph50	0.000	0.000
Ph59	0.000	0.013
Ph61	0.000	0.000
Ph62	0.050	0.039
Ph70	0.000	0.000
Ph81	0.000	0.000
Ph83_11	0.038	0.000
Ph128	0.035	0.000
Ph142_8	0.016	0.000

Table S3 – Parameter estimates for the clines using HZAR software. Each locus has the respective best fit model, the centre of the cline measure as the distance (meters) from the northernmost sample; width (1/maximum slope); pMin is the frequency of the west end of the cline; pMax frequency at the eastern end of the cline.

Locus	Model	Centre	Width	pMin	pMax
Ph14	Free None	652	0.115747	0.118047	0.804294
Ph21	Free None	650.9846	0.009987	0.11195	0.84521
Ph22	None right	645.0679	319.2657	-	-
Ph31	Free None	631.4199	134.5451	0.146605	0.784251
Ph38	Free None	602.0505	12.52942	0.216449	0.800088
Ph39	Free None	652.7395	1.291378	0.099163	0.804714
Ph43	Free None	652.3998	2.121211	0.085595	0.836935
Ph62	None mirror	617.2438	591.4236	0.002609	0.083664
Ph81	Free None	652.3114	0.494882	0.0159	0.771951
Ph83	None right	650.2788	118.6724	-	-
Ph25	Free None	812.3722	44.06699	0.419112	0.806844
Ph30	Free None	664.045	101.3274	0.025693	0.719161
Ph32	Free None	665.04	105.3277	0.026651	0.739597
Ph61	Free None	582.8305	222.7866	9.326679-5	0.845926
Ph70	Free None	630.6948	67.99728	2.915333-5	0.770743
Ph128	null model	-	-	-	-
mtDNA	Fixed right	646.6031	124.843	0	1
Structure	Free None	642.9061	96.23818	0.004318	0.89153